Long Term Evolution-Advanced and Future Machine-to-Machine Communication

submitted to the
Faculty of Physics and Electrical Engineering,
University of Bremen

for obtainment of the academic degree

Doktor-Ingenieur (Dr.-Ing.)

Dissertation

by

Safdar Nawaz Khan Marwat, M.Sc. B.Sc.

from Lakki Marwat, Pakistan

First assessor:	Prof. Dr. rer. nat. habil. Carmelita Görg
Second assessor:	Prof. Dr.-Ing. Ralf Lehnert
Submission date:	1 November 2014
Colloquium date:	12 December 2014

I assure, that this work has been done solely by me without any further help from others except for the official support of the Communication Networks group of the University of Bremen. The literature used is listed completely in the bibliography.

Bremen, 1st of November 2014

(Safdar Nawaz Khan Marwat)

Dedication

I dedicate this doctoral research thesis to my parents who lived far away from me during my studies but provided me with their prayers and support for which I am extremely grateful to them. I would like to thank my wife for her patience and unequivocal support during these tough academic years. I express my gratitude to my kids, who had to sacrifice their recreational activities and playtime due to my academic commitments. I am greatly indebted to my family and relatives who provided me encouragement and inspiration in achieving my goals. Last but not the least, I dedicate this work also to my siblings and especially to my younger brother who was engaged in defending the frontiers of the motherland as an army officer during my doctoral studies.

Acknowledgments

I would like to acknowledge the support, motivation and encouragement of all those cordial people who helped me in writing this thesis. I wish that I could mention all of them but I resort to name only a few. This thesis would not have been accomplished without the help, support and patience of Prof. Dr. rer. nat. habil. Carmelita Görg, whose knowledge, encouragement, motivation and guidance have been of immense worth to me. I would like to express my gratefulness to Prof. Dr.-Ing. Andreas Timm-Giel for his kindness and cooperation.

I am greatly indebted to the support and collaboration of Dr. Yasir Zaki, Dr. Xi Li and Dr. Thushara Weerawardane. They have played an exceptional role in teaching and mentoring me. I extend my warmest thanks to Dr. Umar Toseef, Asanga Udugama and Dr. Koojana Kuladinithi for providing me great support, and Thomas Pötsch for being a great work partner throughout all the activities we performed together. I would like to thank Dr. Andreas Könsgen and Dr. Markus Becker for offering their expertise, lots of gratitude to Dr. Mohammad Muttakin Siddique, Amanpreet Singh, Dr. Bernd-Ludwig Wenning, Yasir Mehmood, Martina Kammann and Liang Zhao for their support and encouragement during these tough years. I would like to particularly mention Martina Kammann for providing help in official tasks and the affectionate Karl-Heinz Volk who helped in sorting out issues with servers, software licenses and hardware malfunctions. I feel obliged to acknowledge the help of my colleagues of the Communication Networks group at the Hamburg University of Technology, especially Ming Li for his simulation proficiency. I would also like to thank all my students for their dedication and good work.

I express my gratitude to the University of Engineering and Technology, Peshawar, Pakistan, for providing me the opportunity to obtain a Master degree and finish my doctoral studies in Germany. In the end, I acknowledge the role of the International Graduate School for Dynamics in Logistics, University of Bremen for providing a structured doctoral training, supporting my publications and financing business travels. I am obliged to Dr. Ingrid Rügge, Karolin Halmai-Samel, Almut Drüner and all IGS graduates who helped me throughout my doctoral research.

Safdar Nawaz Khan Marwat

Abstract

Long Term Evolution (LTE) has adopted Orthogonal Frequency Division Multiple Access (OFDMA) and Single Carrier Frequency Division Multiple Access (SC-FDMA) as the downlink and uplink transmission schemes respectively. Quality of Service (QoS) provisioning is one of the primary objectives of wireless network operators. In LTE-Advanced (LTE-A), several additional new features such as Carrier Aggregation (CA) and Relay Nodes (RNs) have been introduced by the 3rd Generation Partnership Project (3GPP). These features have been designed to deal with the ever increasing demands for higher data rates and spectral efficiency. The RN is a low power and low cost device designed for extending the coverage and enhancing spectral efficiency, especially at the cell edge.

Wireless networks are facing a new challenge emerging on the horizon, the expected surge of the Machine-to-Machine (M2M) traffic in cellular and mobile networks. The costs and sizes of the M2M devices with integrated sensors, network interfaces and enhanced power capabilities have decreased significantly in recent years. Therefore, it is anticipated that M2M devices might outnumber conventional mobile devices in the near future.

3GPP standards like LTE-A have primarily been developed for broadband data services with mobility support. However, M2M applications are mostly based on narrowband traffic. These standards may not achieve overall spectrum and cost efficiency if they are utilized for serving the M2M applications.

The main goal of this thesis is to take the advantage of the low cost, low power and small size of RNs for integrating M2M traffic into LTE-A networks. A new RN design is presented for aggregating and multiplexing M2M traffic at the RN before transmission over the air interface (Un interface) to the base station called eNodeB. The data packets of the M2M devices are sent to the RN over the Uu interface. Packets from different devices are aggregated at the Packet Data Convergence Protocol (PDCP) layer of the Donor eNodeB (DeNB) into a single large IP packet instead of several small IP packets. Therefore, the amount of overhead data can be significantly reduced.

The proposed concept has been developed in the LTE-A network simulator to illustrate the benefits and advantages of the M2M traffic aggregation and multi-

plexing at the RN. The potential gains of RNs such as coverage enhancement, multiplexing gain, end-to-end delay performance etc. are illustrated with help of simulation results.

The results indicate that the proposed concept improves the performance of the LTE-A network with M2M traffic. The adverse impact of M2M traffic on regular LTE-A traffic such as voice and file transfer is minimized. Furthermore, the cell edge throughput and QoS performance are enhanced. Moreover, the results are validated with the help of an analytical model.

Kurzfassung

LTE (Long Term Evolution) nutzt OFDMA (Orthogonal Frequency Division Multiple Access) und SC-FDMA (Single Carrier Frequency Division Multiple Access) als Downlink- bzw. Uplink-Zugangsverfahren. Die Einhaltung der Dienstgüte (QoS, Quality of Service) ist dabei eines der primären Ziele der Mobilfunkanbieter. In LTE-A (LTE-Advanced) wurden einige zusätzliche neue Merkmale wie Carrier Aggregation (CA) und Relay Nodes (RNs) von der 3GPP (3rd Generation Partnership Project) eingeführt. Diese Funktionen wurden entwickelt, um den ständig steigenden Anforderungen an höheren Datenraten und spektraler Effizienz zu begegnen. Der RN ist ein low-power und low-cost Gerät zur Erweiterung der Funkabdeckung und der spektralen Effizienz, vor allem an den Rändern der Zellen.

Mobile Netze müssen sich künftigen Herausforderungen stellen. Dazu gehört der erwartete Anstieg von Machine-to-Machine (M2M) Verkehr in zellulären mobilen Netzen. Kosten und Größe von M2M-Geräten mit integrierten Sensoren, Netzschnittstellen und verbessertem Leistungsverbrauch sind in den letzten Jahren erheblich gesunken. Daher wird davon ausgegangen, dass die Anzahl von M2M-Geräten in naher Zukunft die Anzahl konventioneller mobiler Geräte übertreffen wird.

3GPP Standards wie LTE-A wurden in erster Linie für Breitband-Datendienste mit Unterstützung von Mobilität entwickelt. Einzelne M2M-Anwendungen basieren meist auf Schmalband-Verkehr, aber sie benötigen ggf. einen erheblichen Teil der Ressourcen aufgrund der Anzahl der Geräte. Die bisherigen Standards können insgesamt keine Spektrums- und Kosteneffizienz erreichen, wenn sie für M2M-Anwendungen genutzt werden.

Das Hauptziel der vorliegenden Arbeit ist die Integration der kostengünstigen und effizienten RNs für M2M-Verkehr in LTE-A-Netzen. Ein neues RN-Design wird präsentiert für die Aggregation und das Multiplexen von M2M-Verkehr in dem RN vor der Übertragung über die Luftschnittstelle (Un-Schnittstelle) an die Basisstation (eNodeB). Die Datenpakete der M2M-Geräte werden über die Uu-Schnittstellen an den RN gesendet. Mehrere kleine IP-Pakete von verschiedenen Geräten werden zu einem einzigen großen IP-Paket auf der PDCP (Packet Data Convergence Protocol)-Schicht zusammengefasst. Daher kann die zu übertragen-

de Overhead-Datenmenge deutlich reduziert werden.

Das vorgeschlagene Konzept wurde in einem LTE-A-Netzsimulator implementiert, um die Vorteile der M2M-Verkehr-Aggregation und des Multiplexens in dem RN zu veranschaulichen. Die potenziellen Vorteile der RNs wie etwa Verbesserung der Abdeckung, Multiplex-Gewinn, Verhalten der Ende-zu-Ende-Verzögerung, etc., werden mit Hilfe von Simulationsergebnissen illustriert.

Die Ergebnisse zeigen, dass das vorgeschlagene Konzept die Leistung des LTE-A-Netzes mit M2M-Verkehr verbessert. Negative Auswirkungen des M2M-Verkehrs auf regulären LTE-A-Datenverkehr wie Sprach- und Datei-Transfer werden minimiert. Darüber hinaus wird der Durchsatz an den Rändern der Zelle und die Dienstgüte (QoS) verbessert. Darüber hinaus werden die Ergebnisse mit Hilfe eines analytischen Modells überprüft.

Contents

List of Figures

List of Tables

List of Abbreviations

1G	first generation
2G	second generation
3G	third generation
4G	fourth generation
3GPP	3rd Generation Partnership Project
AAL2	ATM Adaptation Layer 2
AC	Admission Control
ACK	ACKnowledgement
aGW	access GateWay
AMC	Adaptive Modulation and Coding
AMPS	Advanced Mobile Phone System
ARP	Allocation and Retention Priority
ARQ	Automatic Repeat reQuest
ATB	Adaptive Transmission Bandwidth
ATM	Asynchronous Transfer Mode
AVI	Average Value Interface
BCH	Broadcast CHannel
BET	Blind Equal Throughput
BQA	Bandwidth and QoS Aware
BSR	Buffer Status Report
CA	Carrier Aggregation
CC	Component Carrier
CCS	Component Carrier Selection
CDMA	Code Division Multiple Access
CFI	Control Format Indicator
CLPC	Closed Loop Power Control
CoMP	Coordinated MultiPoint
CPS	Common Part Sub-layer
CQA	Channel and QoS Aware
CQI	Channel Quality Indicator
CSI	Channel State Information
DCI	Downlink Control Information
DeNB	Donor eNodeB
DFTS	Discrete Fourier Transform-Spread
DSCH	Downlink Shared CHannel
EESM	Effective Exponential SINR Mapping
EMA	Exponential Moving Average
EPC	Evolved Packet Core
EPS	Evolved Packet System
ETSI	European Telecommunications Standards Institute
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FDPS	Frequency Domain Packet Scheduling
FFT	Fast Fourier Transform

FPC	Fractional Power Control
FSCFB	Fixed Size Chunk and Flexible Bandwidth
FTP	File Transfer Protocol
GBR	Guaranteed Bit Rate
G-PFS	GBR-aware Proportional Fair Scheduled
GSM	Global System for Mobile Communication
GTP	GPRS Tunneling Protocol
HARQ	Hybrid Automatic Repeat reQuest
HSDPA	High Speed Downlink Packet Access
HTTP	HyperText Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
IPC	Interference based Power Control
IMS	IP Multimedia Sub-System
ITS	Intelligent Traffic Systems
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
LA	Link Adaptation
LST	Laplace-Stieltjes Transform
LTE	Long Term Evolution
LTE-A	LTE-Advanced
MAC	Medium Access Control
M2M	Machine-to-Machine
MBMS	Multimedia Broadcast Multicast Service
MBR	Maximum Bit Rate
MBSFN	Multimedia Broadcast Single Frequency Network
MCH	Multicast CHannel
MCS	Modulation and Coding Scheme
MIMO	multiple-input and multiple-output
MME	Mobility Management Entity
MT	Maximum Throughput
NACK	Non-ACKnowledgement
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OLPC	Open Loop Power Control
PAPR	Peak-to-Average-Power-Ratio
PBCH	Physical Broadcast CHannel
PBR	Prioritized Bit Rate
PC	Power Control
PCC	Primary Component Carrier
PCFICH	Physical Control Format Indicator CHannel
PCH	Paging CHannel
PDCCH	Physical Downlink Control CHannel
PDCP	Packet Data Convergence Protocol
PDF	Probability Distribution Function
PDSCH	Physical Downlink Shared CHannel
PDN-GW	Packet Data Network GateWay
PDU	Protocol Data Unit
PF	Proportionally Fair
PFS	Proportional Fair Scheduled

PHICH	Physical HARQ Indicator CHannel
PMCH	Physical Multicast CHannel
PRACH	Physical Random Access CHannel
PRB	Physical Resource Block
PS	Packet Scheduling
PSD	Power Spectral Density
PUCCH	Physical Uplink Control CHannel
PUSCH	Physical Uplink Shared CHannel
QAM	Quadrature Amplitude Modulation
QCI	QoS Class Identifier
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RACH	Random Access CHannel
RAT	Radio Access Technology
RC	Resource Chunk
RLC	Radio Link Control
RN	Relay Node
RRC	Radio Resource Control
RRM	Radio Resource Management
RSRP	Received Signal Received Power
SCC	Secondary Component Carrier
SC-FDMA	Single Carrier Frequency Division Multiple Access
S-GW	Serving Gateway
SINR	Signal to Interference and Noise Ratio
SNR	Signal to Noise Ratio
SRS	Sounding Reference Signal
TDPS	Time Domain Packet Scheduling
TIA	Telecommunications Industry Association
TTI	Transmission Time Interval
uDFS	unique Depth-First Search
UDP	User Datagram Protocol
UE	User Equipment
UMTS	Universal Mobile Telecommunication System
USA	United States of America
USCH	Uplink Shared CHannel
USIM	Universal Subscriber Identity Module
VoIP	Voice over Internet Protocol
WAN	Wide Area Networks
WCDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
W-PF	Weighted Proportional Fair
W-PFS	Weighted Proportional Fair Scheduled

List of Symbols

Symbol	Meaning
P	uplink transmit power [dBm]
$P_{\max}$	maximum allowed uplink transmit power [dBm]
P_0	uplink transmit power control parameter
α	path loss compensation factor
L	path loss [dBm]
Δ_{MCS}	uplink power control parameter for upper layers
Δ_i	closed-loop correction value
N	number of PRBs allocated to a UE under PC
β	interference factor
L_{other}	logarithmic sum of path losses towards other cells [dBm]
M	number of mobile users
$M_{\max}$	maximum number of users scheduled
H	number of users with pending HARQ
R_{km}	distance [km]
S	normally distributed shadowing [dB]
μ_{dB}	mean of S [dB]
σ_{dB}	standard deviation of S [dB]
δ	distance [m]
$A(\delta)$	shadowing covariance between points at a distance δ apart
d_c	decorrelation distance [m]
${\sigma_{\mathrm{dB}}}^2$	maximum value of $A(\delta)$
V_i	independent identically distributed normal random variables
$\Lambda_i(t)$	TDPS metric value of user i
$R_{\mathrm{avg},i}(t)$	average throughput of user i at time t over EMA [Kbps]
T_w	EMA time window length [TTI]
$R_{\mathrm{ach},i}(t)$	actual bit rate of user i in previous TTI [Kbps]
n_i	maximum number of obtainable PRBs for user i under PC
$R_{\mathrm{inst},i}(t, n_i)$	instantaneously achievable wideband throughput of user i at time t with n_i PRBs [Kbps]

Symbol	Meaning
a	bearer index
$W_{i,a}(t)$	QoS weight of the bearer a of user i at time t
$R_{\min,a}$	bit rate budget of QoS class of bearer a [Kbps]
$\tau_{\max,a}$	end-to-end delay budget of QoS class of bearer a [ms]
$R_{\text{avg},i,a}(t)$	average throughput of bearer a [Kbps]
$\tau_{i,a}(t)$	packet delay of bearer a of UE i [ms]
$\rho_a(t)$	variable having value 10 if delay threshold is breached, otherwise 1
$\lambda_{i,c}$	FDPS metric value of user i at PRB c
$SINR_{i,c}(t)$	SINR value of user i at PRB c [dB]
$R_{\text{inst},i,c}$	instantaneously achievable throughput of user i at PRB c [Kbps]
$w_i(t)$	FDPS weighting factor for users below their minimum required bit rates
R_i	Uu throughput of user i [Kbps]
R_{avg}	average Uu throughput of all users [Kbps]
$L_{\text{threshold}}$	path loss threshold [dBm]
$L_{95\%}$	estimated 95 percentile user path loss [dBm]
N_{CC}	total number of CCs
P_{backoff}	estimated power backoff [dBm]
$N_{\max}$	maximum PRBs allocated to RN
$T_{\max}$	maximum waiting time for RN multiplexing [ms]
$n_{\max}$	maximum buffer size for RN multiplexing [bits]
k	number of packets in the multiplexing buffer
r	number of packets multiplexed in a TBS
l	fixed size of packet arriving at PDCP layer of relay node access link
p_k	probability of next arrival before multiplexing
q_k	probability of multiplexing before next arrival
ϕ_k	product of transition probabilities in r-coxian process at k^{th} arrival
C_k	stage counter of aggregation buffer
$N_{k,MCS}$	number of PRBs required in state k depending on MCS
K	discrete random variable representing the number of small packets multiplexed into large packet
$\overline{N}$	average PRBs per multiplexing process

1 Mobile Network Evolution

Wireless communication is one of the most rapidly growing industries. Cellular systems are substituting the wired Wide Area Networks (WAN). Similarly, Wireless Local Area Networks (WLANs) are replacing the wired networks in homes, offices and campuses. However, the growing number of mobile subscribers and emerging applications are posing technical challenges in terms of Quality of Service (QoS), performance and capacity.

Historically, wireless communication started by using light, sound, smoke and flag as wireless signals. Inventions like telegraph and telephone in the 19th century replaced such forms of communication. In 1895, Guglielmo Marconi demonstrated 1st ever radio transmission. In 1901, he performed transatlantic radio transmissions and later, in 1920, he discovered short waves which enabled long distance radio transmissions.

The initial systems of mobile communication were deployed in the United States in the 1940s and in Europe in the 1950s. Considering the growing demand of wireless service, the Advanced Mobile Phone System (AMPS) was introduced in 1983 as an analogue system. It is known as the first generation (1G) mobile phone system. Meanwhile, Groupe Spécial Mobile (GSM), which was formed in 1982, took the task of developing a digital mobile system to be mainly used in Europe. In 1990, phase I of the GSM specifications was finalized. This standard is known today as the Global System for Mobile Communication (GSM) or second generation (2G) system. In 1991, the 1st GSM network was deployed in Finland. In 1996, the Universal Mobile Telecommunication System (UMTS) forum was formed in Zurich. The 1st UMTS network, which is also known as the third generation (3G) system, was launched in the United Kingdom in 2003. The 3rd Generation Partnership Project (3GPP) is responsible for defining the system specifications for 3G systems.

1.1 Long Term Evolution

Under 3GPP, a workshop was organized in Toronto, Canada for the evolution of 3G systems in November 2004 [Sho04]. More than 40 contributions were presented in

the workshop which highlighted proposals and ideas. The high-level requirements for the evolved system defined in the workshop were as follows:

- lower cost per bit
- more services at lower costs
- adaptable use of prevailing and new frequency bands
- uncomplicated architecture with open interfaces
- fair power consumption at the terminal

In December 2004, a feasibility study was commenced by 3GPP for the Long Term Evolution (LTE) system. The motivation for this study was to establish a plan of action for the evolution of a packet only network with higher data rates, lower latency providing high QoS and simple infrastructure. The details of the performance requirements can be summarized as follows (taken from [25.09]).

- *Peak Data Rate*: Downlink, 100 Mbps within a 20 MHz downlink spectrum allocation (5 bps/Hz); uplink, 50 Mbps (2.5 bps/Hz) within a 20 MHz uplink spectrum allocation
- *Control Plane Latency*: Transition time below 100 ms from a camped state to an active state; transition time below 50 ms between a dormant state and an active state
- *Control Plane Capacity*: Support for minimum 200 users in a cell in the active state for up to 5 MHz bandwidth allotment
- *User Plane Latency*: Below 5 ms in condition of single user with single data stream for small for IP packet
- *User Throughput*: Downlink: average user throughput per MHz, 3 to 4 times Release 6 High Speed Downlink Packet Access (HSDPA); Uplink: average user throughput per MHz, 2 to 3 times Release 6 Enhanced Uplink
- *Spectrum Efficiency*: Downlink: spectrum efficiency (bits/sec/Hz/site) 3 to 4 times Release 6 HSDPA in loaded condition; Uplink: spectrum efficiency (bits/sec/Hz/site) 2 to 3 times Release 6 Enhanced Uplink in loaded condition

- *Mobility*: Network optimized for mobile speed of 0 to 15 km/h; High performance support for mobile speed between 15 and 120 km/h; Communication in network to be sustained at speeds of 120 km/h to 350 km/h (or even up to 500 km/h in certain frequency bands)
- *Coverage*: Requirements regarding throughput, spectrum efficiency and mobility to be met for 5 km cells; with minor deterioration for 30 km cells; even cell ranges up to 100 km to be acceptable
- *Multimedia Broadcast Multicast Service*: Reduced terminal complexity in terms of modulation, coding and User Equipment (UE) bandwidth along with multimedia service
- *Spectrum Flexibility*: Ability of radio network to operate in various sizes of spectrum allocations ranging from 1.25 MHz to 20 MHz in both the uplink and downlink along with the ability to operate in paired and unpaired spectrum
- *Co-existence and Inter-working with other 3GPP Radio Access Networks*: Ability of the LTE radio network to operate simultaneously with GSM and UMTS radio networks in a particular location with interruption time less than 300 ms in case of inter-system handover
- *Architecture*: Packet based architecture to be designed supporting the end-to-end QoS (e.g. to guarantee customers that end-to-end latency remains under a specific level)
- *Complexity*: Reduce the number of alternatives and eradicate repetitious mandatory features.

The accomplishment of the wide-ranging requirements summarized above demands the employment of advanced technology [STB09]. The recent progress in mobile radio technology has made the task of realizing these requirements achievable. The technology proposed for LTE include features such as the multicarrier technology, the multiple-antenna technology, and the utilization of packet-switching for the radio interface.

1.1.1 Architecture

LTE is the 1st mobile communication system designed right from the beginning to be a packet-switched network instead of being circuit-switched like its predecessor networks. The LTE architecture provides no support for circuit switching.

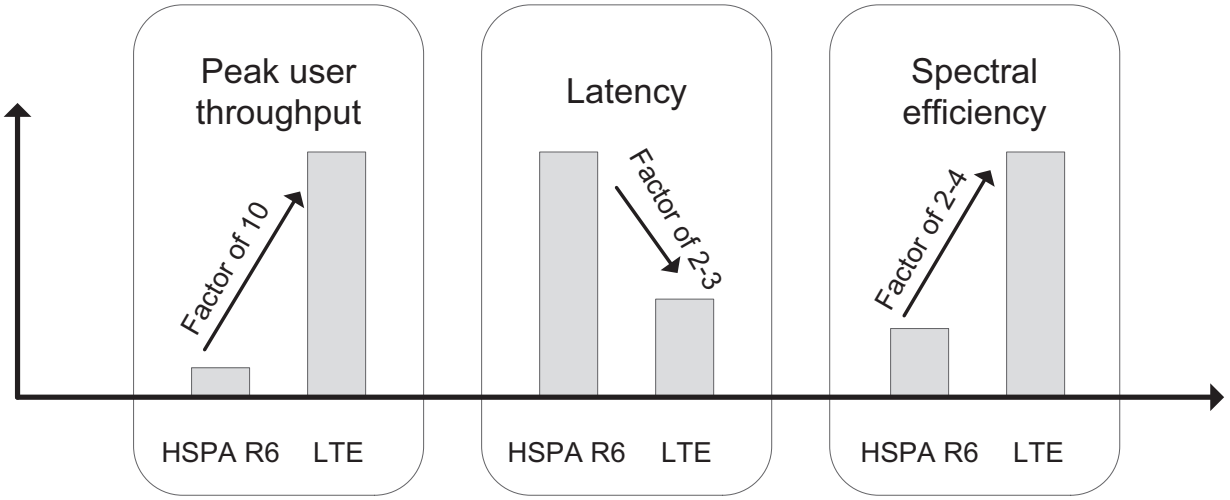

Figure 1.1: Main performance targets of LTE [HT09]

All the layers of the LTE protocol stack (discussed in subsection 1.1.2) have been designed under the packet switching paradigm. LTE has a flat architecture design which consists of only one type of node, the base station in the radio access network. The term flat architecture means that there is no central controller for normal and non-broadcast users. Protocols between UE and base station are called Access Stratum protocols. In LTE, the base station is known as the evolved NodeB or eNodeB. UMTS and GSM had other nodes in addition to the base station in the radio access networks; Radio Network Controller in UMTS and Base Station Controller in GSM. LTE is also designed to support easy deployment and configuration.

The LTE architecture can be divided into four domains (Figure 1.2): UE, Evolved Universal Terrestrial Radio Access Network (E-UTRAN), Evolved Packet Core (EPC) and Services domain [HT09]. The X2 interface interconnects the eNodeBs, while the S1 interface connects eNodeB and EPC. The three domains; UE, E-UTRAN and EPC are termed as Evolved Packet System (EPS). The role of EPS is the provision of Internet Protocol (IP) based connection services. The E-UTRAN and EPC are packet networks. Absence of the circuit switched domain from the core network is a major architectural change in LTE. The IP Multimedia Sub-System (IMS) and the Internet are examples of external networks in the service domain. The components of the LTE system architecture are discussed in brief next.

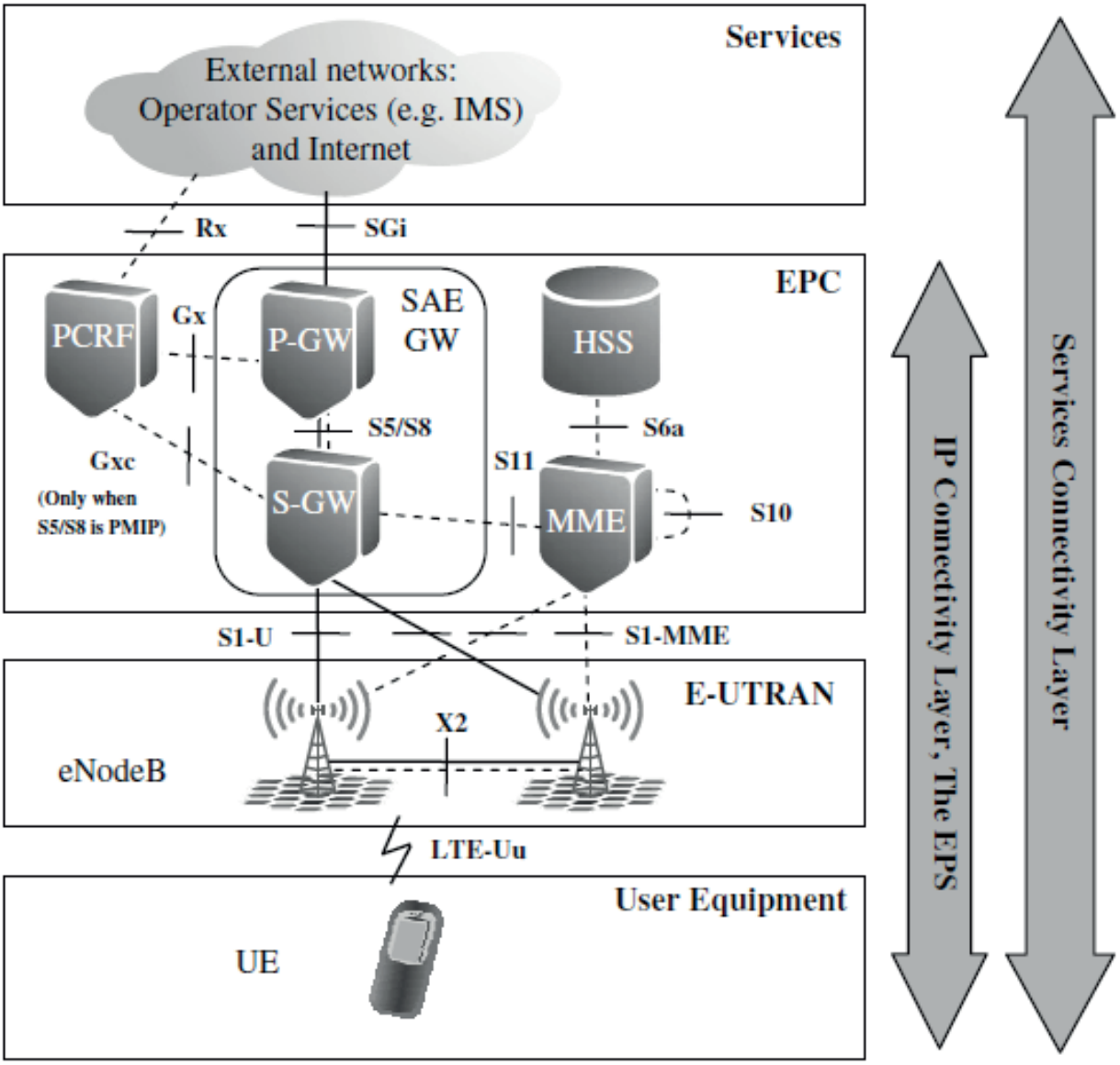

Figure 1.2: LTE system architecture [HT09]

1.1.1.1 User Equipment

A UE is a hand held end user device for communication. It contains another module called Universal Subscriber Identity Module (USIM) used for user identification and authentication. The UE performs signaling with the network to set up, maintain and terminate the communication links requested by the end user. The UE is an interface provider to the end user for utilization of various applications; such as the Voice over Internet Protocol (VoIP) etc.

1.1.1.2 E-NodeB

The eNodeB is the main node in the E-UTRAN. The eNodeBs are installed all over the coverage area of the network. All the functionalities related to radio, in the fixed part of the LTE system, are handled at the eNodeB. Radio Resource Management (RRM) is among the basic tasks of eNodeB. Examples of RRM funtionalities include radio resource allocation to user requests with QoS provision. The eNodeB also performs tasks related to mobility management such as handover decisions based on signal power measurements, since each UE can be served by only a single eNodeB.

1.1.1.3 Mobility Management Entity

The Mobility Management Entity (MME) is the primary control element of the EPC. MME functionalities are related only to the control plane and have no involvement in the user plane path. The MME provides authentication service when a UE enters the network. After authentication, the UE is provided with security from eavesdropping (call tapping) and unauthorized UE tracking by MME. The MME facilitates mobility management by continuously tracking the location of all the UEs in its service area. The MME also manages the subscription profile of the UEs. The subscription data of a UE registering to the network is acquired from its home network. A request for a service by a UE that is not supported in its profile would be rejected by the MME.

1.1.1.4 Serving Gateway

The Serving Gateway (S-GW) is the node which facilitates the transfer of IP packets. During UE movements between eNodeBs, the S-GW acts as a local mobility anchor. Downlink data received at the S-GW is buffered if the UE is idle. The MME has to perform the task of UE paging to establish the bearer so that the S-GW can release the buffered data. The S-GW also collects user charging information and can facilitate interception by law enforcing agencies. It can also serve as mobility anchor for working with other 3GPP technologies.

1.1.1.5 Packet Data Network Gateway

The Packet Data Network GateWay (PDN-GW) performs the task of allocating IP addresses to the UEs. The PDN-GW ensures the provision of QoS to users and that the users are charged according to the charging rules. The downlink IP packets are

filtered according to bearers based on QoS. It acts as a mobility anchor for working with non-3GPP technologies like Worldwide Interoperability for Microwave Access (WiMAX).

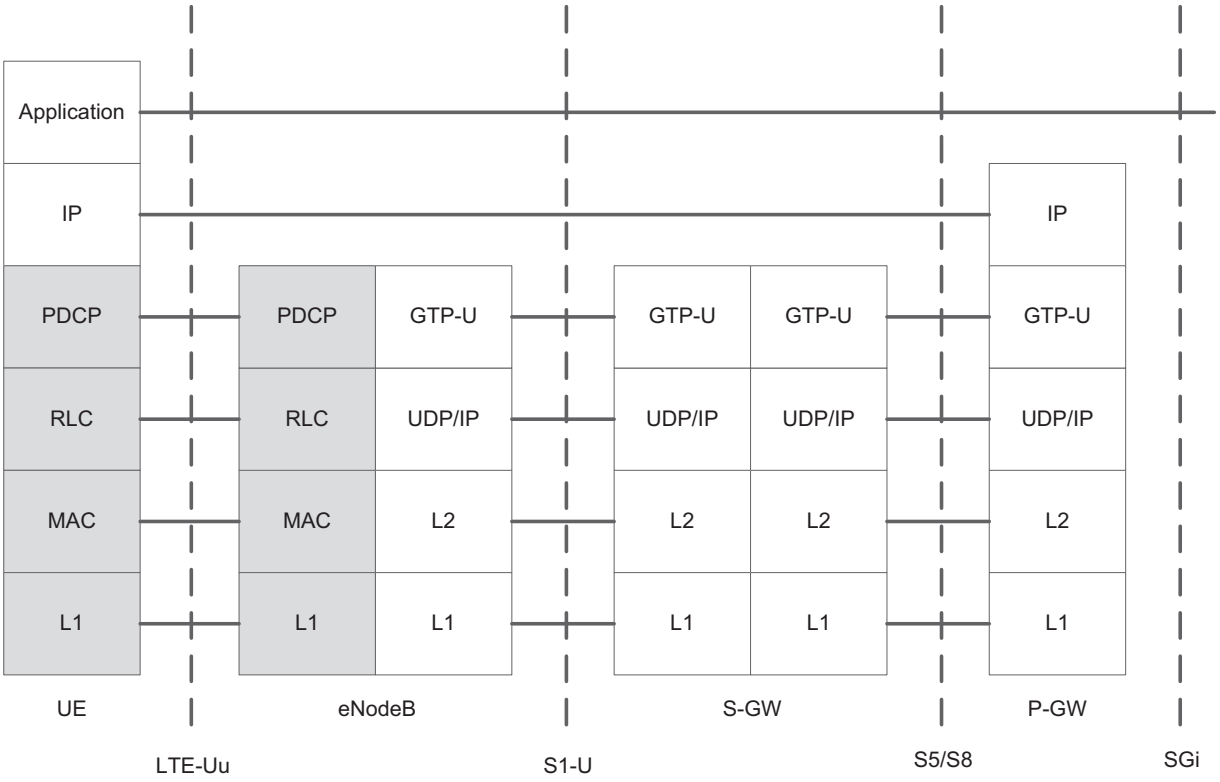

Figure 1.3: LTE user plane protocol stack

1.1.2 Protocol Stack

The Protocol stack of LTE is composed of two planes; the user plane and the control plane. The user plane protocol stack is composed of protocols within the UE, eNodeB, S-GW and PDN-GW as in Figure 1.3. The control plane protocol stack is composed of protocols in the UE, the eNodeB and the MME as in Figure 1.4. The greyed protocols correspond to the E-UTRAN protocol stack and consist of Radio Resource Control (RRC), Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC) and Medium Access Control (MAC) and Physical Layer (or L1).

- The RRC (only in control plane) controls the utilization of radio resources and facilitates handover. System broadcast information handling is a task of RRC. It also handles the UE's signaling and data connections. Inter Radio Access Technology (RAT) mobility is also facilitated by the RRC.

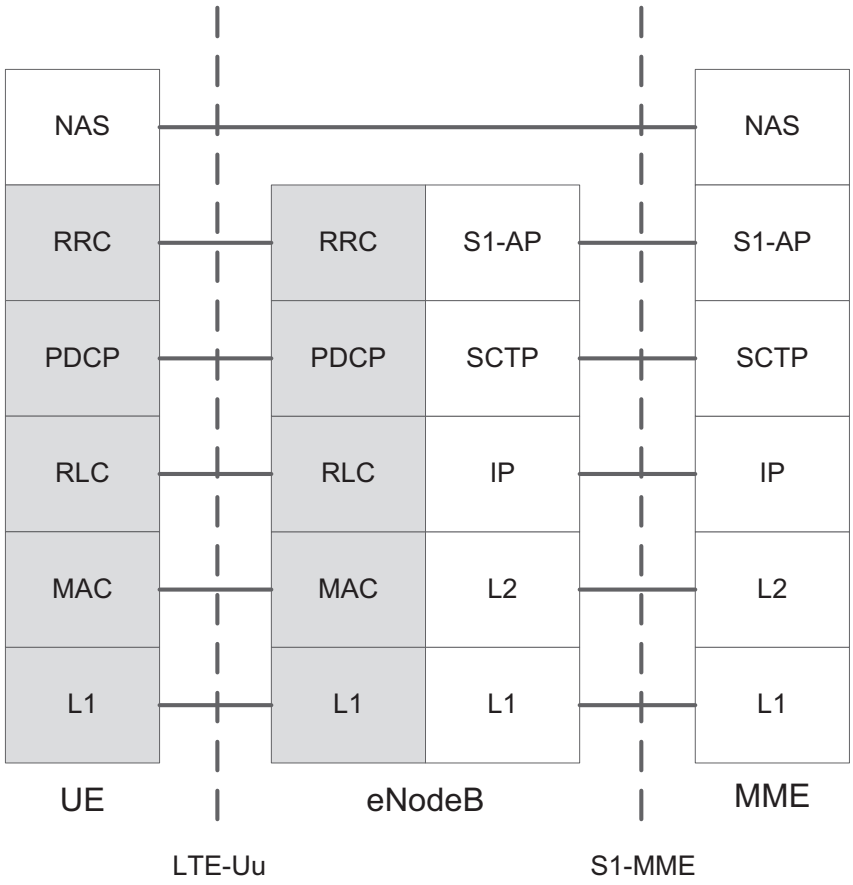

Figure 1.4: LTE control plane protocol stack

- The PDCP performs the tasks of IP header compression in user plane, while encryption and integrity protection in control plane. It also supports reordering and retransmission of upper layer Protocol Data Units (PDUs) during handover.

- The RLC performs segmentation and concatenation of the upper layer PDUs to tailor them for radio interface transmission. Error correction is achieved at the RLC layer by using the Automatic Repeat reQuest (ARQ) method. Ordering of packets received out of order is done at the RLC layer.

- The MAC layer performs dynamic data scheduling according to priorities, and multiplexing of data to physical layer transport blocks. The MAC directs the RLC about the amount of data each bearer can transmit so that the RLC can sent data packets accordingly. Error correction is also performed using Hybrid ARQ (HARQ).

- L1 is the physical layer of the E-UTRAN radio interface and deals with radio access schemes, channel fading, transmit power etc.

1.1.3 Air Interface

Various schemes had been proposed for the selection of the LTE air interface scheme during the initial phases of LTE standard development. A key decision made in this regard was to employ a multicarrier approach for multiple-access in LTE. Initially, a set of candidate schemes were proposed. The downlink candidate schemes were Orthogonal Frequency Division Multiple Access (OFDMA) and Multiple Wideband Code Division Multiple Access (WCDMA). The uplink candidate schemes were Single Carrier Frequency Division Multiple Access (SC-FDMA), OFDMA and Multiple WCDMA. In December 2005, the decision of choosing multiple-access schemes was announced. OFDMA was chosen for downlink and SC-FDMA for uplink. These schemes provide flexibility in the frequency domain as explained in the following subsections. The choice of these schemes was facilitated by the availability of their transceivers at reasonable prices.

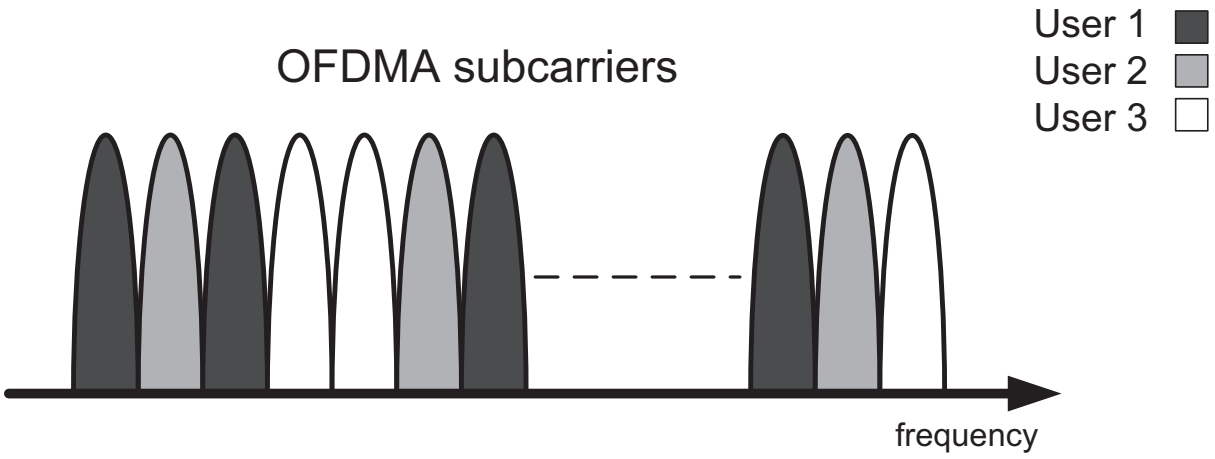

Figure 1.5: OFDMA uplink frequency domain view

1.1.3.1 Orthogonal Frequency Division Multiple Access

OFDMA is the multiuser extension of OFDM and underlying technology. OFDM splits the spectrum into smaller subcarriers closely spaced and orthogonal to each other for the transmission of a single data stream using existing modulation schemes like Quadrature Phase Shift Keying (QPSK) and 16-Quadrature Amplitude Modulation (16-QAM). OFDMA utilizes the spectrum division of OFDM by allocating

subsets of subcarriers to individual users [STB09]. It is possible to allocate the resources of variable bandwidth to different users (Figure 1.5). Therefore, users can be freely scheduled in the frequency domain. OFDMA also allows a flexible use of spectrum by utilizing different bandwidths without changing the basic system parameters and design of equipment. This is called bandwidth scalability.

The traditional Frequency Division Multiplexing (FDM) techniques waste the spectrum. The parallel modulated carriers should be prevented from interfering with each other. This can be accomplished by spacing them with guard bands resulting in bandwidth inefficiency. Additionally, receiving filters should be capable of separating the parallel carriers from each other. In traditional FDM systems, a deep fade in power level or an interferer can be sufficient to terminate the link between transmitter and receiver.

In OFDM systems, the carriers can be tightly spaced and still easily separated from each other. This results in efficient usage of the frequency spectrum. In case of deep fade or strong interference, only a few of the subcarriers might be disturbed allowing transmission to continue through the less affected subcarriers.

Drawbacks of the OFDMA include the expensive transmitter design [STB09]. OFDMA signals have relatively higher Peak-to-Average Power Ratio (PAPR). Therefore, highly linear radio frequency power amplifiers are required, resulting in power inefficiency. But in case of LTE downlink transmissions, OFDMA is feasible because high-cost implementation and high power requirements for the base station are not significant issues. However, these issues would be significant in case of the mobile terminal in the uplink transmission.

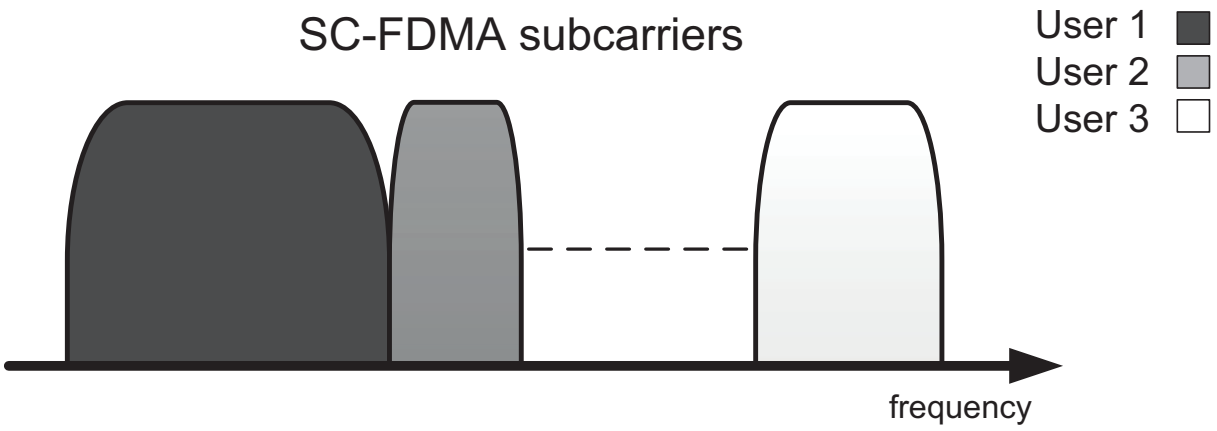

Figure 1.6: SC-FDMA uplink frequency domain view

1.1.3.2 Single Carrier Frequency Division Multiple Access

SC-FDMA is also an OFDM based multiple-access scheme. It has been chosen for LTE uplink over OFDMA because of a relatively lower Peak-to-Average-Power-Ratio (PAPR) of SC-FDMA signals. In uplink, the mobile terminal power is a high priority concern and is addressed by the selection of SC-FDMA as the uplink radio interface scheme. This results in increasing the battery life of the mobile terminal. SC-FDMA also provides frequency domain flexibility. It is possible to allocate the resources of variable bandwidth to different users. Also, utilization of different spectrum bandwidths without changing the basic system parameters and design of equipments is possible with SC-FDMA. A major constraint regarding SC-FDMA is that the subcarriers allocated to a single mobile terminal should be adjacent, i.e. contiguous to each other (Figure 1.6). Thus, the design of uplink scheduling algorithms is quite challenging due to this constraint.

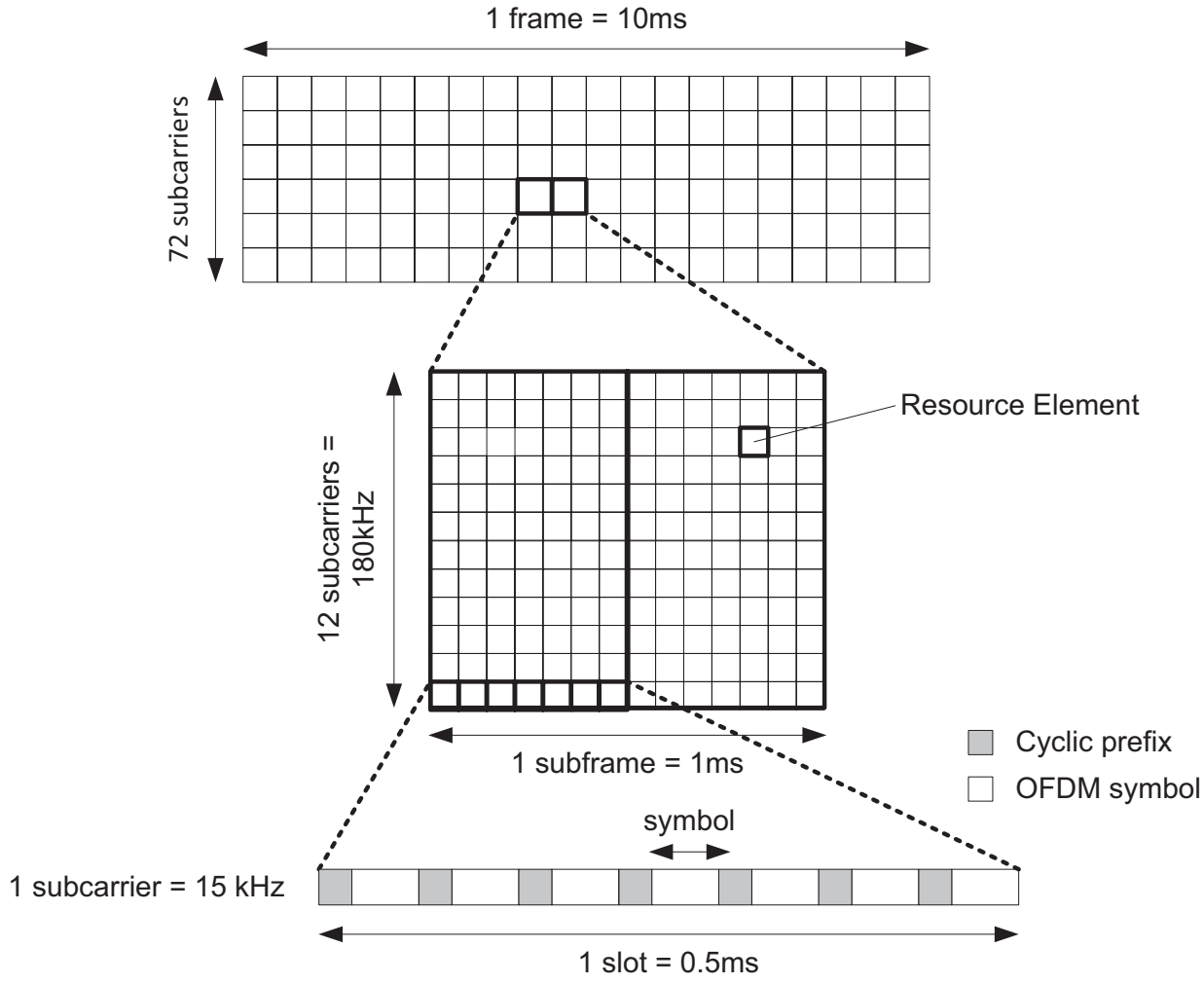

Figure 1.7: LTE frame structure

1.1.3.3 LTE Frame Structure

The resources in LTE are divided into time and frequency domain units [STB09]. In time domain, a frame has the length of 10ms. The frame is divided into subframes of length 1ms. Each subframe consists of two slots of 0.5ms and each slot has seven OFDM symbols if a normal length of cyclic prefix is used (six symbols for extended cyclic prefix). In frequency domain, a bandwidth of 15kHz is occupied by a subcarrier. A group of 12 subcarriers (180kHz) having a length of one slot in time domain is called a Physical Resource Block (PRB). The smallest time-frequency resource unit is a subcarrier with the duration of one OFDM symbol, called Resource Element (RE). This frame structure explained here can be easily conceptualized by viewing Figure 1.7.

Available bandwidth (MHz)	Number of PRBs	Number of subcarriers	FFT[1] size
1.4	6	72	128
3	15	180	256
5	25	300	512
10	50	600	1024
15	75	900	1536
20	100	1200	2048

Table 1.1: LTE channel bandwidth

The reason for defining only a limited number of channel bandwidths is to limit the deployment complexity. The bandwidths of 1.4 and 3MhZ are specified only to facilitate GSM and CDMA2000[2] migration to LTE. The LTE channel bandwidth is characterized by the maximum number of PRBs in the channel [HT09]. The bandwidth flexibility feature of LTE ensures that the spectrum should be scalable [DPSB08]. The channel bandwidths specified for LTE are given in Table 1.1.

1.1.4 System Deployment

The deployment of the LTE system is already underway worldwide and various leading telecommunication companies have launched or committed to launch the LTE services [Lte13]. Several cellular service providers have started operation after performing test trial in different countries of the world in collaboration with various radio technology providing companies (Table 1.2).

[1] Fast Fourier Transform
[2] Code Division Multiple Access 2000

Operator	Region	Status
Antel	Uruguay	Launched in December 2011 in Punta del Este and Maldonado
AT&T	United States of America (USA)	Launched in September 2011 in Atlanta, Chicago, Dallas, Houston and San Antonio
Bharti Airtel	India	Launched in April 2012
China Mobile	China	Launched in December 2013
Chunghwa Telecom	Taiwan	Deployment scheduled to begin in 2011
Etisalat	Middle East	Launched in 2011
Interdnestrcom	Transnistria	Launched in June 2012
KDDI	Japan	Launched in September 2012
MetroPCS	USA	Launched on 21 September 2010 in Las Vegas
NTT DoCoMo	Japan	Launched in 2010
SK Telecom	Korea	First operator to launch of LTE-A in June 2013
Sky Telecom	Brazil	Launched in December 2011
Spectranet	Nigeria	Launched in 2013
Sprint	USA	Launched in December 2012
STC	Saudi Arabia	Launched in September 2011
T-Mobile USA	USA	Launched in 2013
Telecom Italia	Italy	Launched in November 2012
Telefonica Germany	Germany	Launched in July 2011
TeliaSonera	Sweden, Norway	Launched in 2009, the 1st operator to launch LTE commercially
Telstra	Australia	Launched in 2012
UNE EPM Telecomunicaciones	Columbia	Launched in December 2011
Verizon Wireless	USA	Launched in December 2010
Vodafone Spain	Spain	Launched in May 2013
Vodafone Germany	Germany	Launched in September 2010
Zain	Saudi Arabia	Launched in 2011
Zong	Pakistan	Acquired spectrum for LTE in April 2014

Table 1.2: LTE operators worldwide

1.1.5 Quality of Service

The term ‘QoS’ describes the treatment expected by a user in a network and encapsulates all the related technologies, protocols and architecture [AY11]. Providing the users with end-to-end QoS is the ultimate goal of network operators. For this

purpose, a uniform treatment to the packet flow across the network has to be guaranteed.

In LTE, the term 'bearer' is the IP packet flow with defined QoS between the PDN-GW and the UE [STB09]. Each EPS bearer in the LTE system is associated with certain QoS parameters. Packets of one bearer type are expected to undergo the same treatment in terms of QoS. Establishment of a new bearer can be initiated by the UE or by the PDN-GW. A default EPS bearer is established when a user is connected to the Packet Data Network. The default bearer provides IP connectivity to the user as long as the user is connected to the PDN-GW. Other bearers established to the PDN-GW are known as dedicated bearers. A dedicated bearer is established when the default bearer is not enough for achieving the required QoS. The QoS parameters of the already established bearers can be modified dynamically. Bearers parallel to the existing ones can also be established in order to facilitate various services simultaneously. In the radio access network, the eNodeB has to ensure that the required QoS for a bearer over the radio interface is provided.

1.1.5.1 Guaranteed Bit Rate Bearers

An EPS bearer with network resources consistently allocated to it is known as a Guaranteed Bit Rate (GBR) bearer. Bit rates higher than GBR may also be allocated to a bearer in case of availability. An example of a GBR bearer is VoIP call. A Maximum Bit Rate (MBR) parameter can also be associated to a GBR bearer if an upper limit for the data rate has to be enforced.

1.1.5.2 Non-Guaranteed Bit Rate Bearers

An EPS bearer is called a Non-GBR bearer if no network resources are consistently allocated to it. Non-GBR does not guarantee any bit rate. Resources are only allocated to Non-GBR if available. Examples of Non-GBR are File Transfer Protocol (FTP) and the HyperText Transfer Protocol (HTTP). The default bearer of the user should be Non-GBR, because it has to be established permanently. A dedicated bearer can be Non-GBR or GBR.

1.1.5.3 QoS Class Identifiers

All radio bearers are associated with a QoS Class Identifier (QCI). A QCI is always distinguished by a bearer priority, bearer packet delay budget (layer 2) and bearer acceptable packet loss rate (layer 2). The experience of a bearer at the eNodeB can be determined by the QCI of that bearer. A set of QCIs has been standardized

and given in Table 1.3 along with their characteristics. The packet scheduler can be expected to prioritize the scheduling of high priority packets over low priority packet.

QCI	Bearer type	Priority	Packet delay budget (ms)	Packet error loss rate	Example services
1	GBR	2	100	10^{-2}	Conversational voice
2	GBR	4	150	10^{-3}	Conversational video
3	GBR	5	300	10^{-6}	Non-conversational video
4	GBR	3	50	10^{-3}	Real Time Gaming
5	NGBR	1	100	10^{-6}	IMS Signaling
6	NGBR	7	100	10^{-3}	Interactive gaming
7	NGBR	6	300	10^{-6}	Live streaming
8	NGBR	8	300	10^{-6}	Buffered streaming
9	NGBR	9	300	10^{-6}	FTP, chat

Table 1.3: LTE QOS Class Identifiers [STB09]

1.1.5.4 Allocation and Retention Priority

Allocation and Retention Priorities (ARPs) are used to decide if the request for a bearer establishment or modification would be accepted or rejected due to resource limitations. Additionally, the eNodeB can also determine with the help of ARP, which bearers to drop if the network faces a severe limitation of resources. ARP has no influence on the kind of treatment a packet of a particular bearer has to face, after successful established of that bearer. Packets of established bearers are treated only in accordance with their QoS requirements.

1.1.5.5 Prioritized Bit Rate

The Prioritized Bit Rate (PBR) is a QoS parameter for UEs with multiple uplink bearers. Such UEs might face a bearer starvation problem if adequate resources are not allocated to low priority bearers. The purpose of the PBR is to provide only a certain configured bit rate to high priority bearers and then provide a bit rate configured for low priority bearers. Each bearer is served according to its priority in such a way that only the configured PBR is provided to the all the bearers in the 1^{st} step. In the next step, resources are allocated to low priority bearers only if high priority bearers have been completely served. The PBR parameter is not related to GBR of a bearer. A Non-GBR bearer can also be configured with a PBR.

The process of serving a UE with multiple uplink bearers can be summarized as follows [36.13].

- Serve the radio bearers in decreasing priority order up to their PBR.
- Serve the radio bearers in decreasing priority order for the remaining resources.

1.1.6 Transport and Physical Channels

The transport channels connect the MAC layer and the physical layers of the E-UTRAN protocol stack by exchanging data between them. Transport channels define how to transmit data over the air interface and what characteristics should the transmitted data have. Physical channels simply correspond to the radio resource elements carrying data that is generated by the upper layers [STB09].

1.1.6.1 Downlink Transport and Physical Channels

The downlink transport channels are as follows:

Broadcast CHannel: The Broadcast CHannel (BCH) works with a pre-defined transport format. It supports broadcasting in the whole cell coverage area. BCH also transports parts of the system information necessary for access to the Downlink Shared CHannel (DSCH).

Downlink Shared CHannel: The DSCH supports Hybrid Automatic Repeat reQuest (HARQ), Adaptive Modulation and Coding (AMC), transmit power variation, dynamic resource allocation. DSCH transports downlink user data and control messages. The parts of system information that are not transported by the BCH are transported on the DSCH.

Paging CHannel: The Paging CHannel (PCH) supports broadcasting to the whole cell coverage area. The PCH transports paging information to UEs. Updates in system information are communicated to the UEs through the PCH.

Multicast CHannel: The Multicast CHannel (MCH) supports Multimedia Broadcast Multicast Service (MBMS) transmission to multiple cells. It also supports broadcast in the whole cell coverage area. The MCH transports user data and control messages that require Multimedia Broadcast Single Frequency Network (MBSFN) where a UE receives and combines synchronized signals from multiple

cells.

The downlink physical channels are as follows:

Physical Broadcast CHannel: The Physical Broadcast CHannel (PBCH) carries basic system information for configuration and operation of other channels in the cell.

Physical Downlink Shared CHannel: The Physical Downlink Shared CHannel (PDSCH) is the main data carrying channel and used for all user data. It is also used for the broadcast of the system information not being carried on the PBCH. Paging messages are also supported since, there is no specific physical layer paging channel in LTE.

Physical Multicast CHannel: The Physical Multicast CHannel (PMCH) is used to support MBMS transmission to multiple cells. It also supports broadcast in the whole cell coverage area.

Physical Control Format Indicator CHannel: The Physical Control Format Indicator CHannel (PCFICH) carries a Control Format Indicator (CFI). CFI specify the number of Orthogonal Frequency Division Multiplexing (OFDM) symbols (usually 1, 2 or 3) used to transmit control channel information for subframe. UE is generally able to obtain the CFI value, without any channel, by trying to decode the control channel with each possible number. However, it would increase the processing load.

Physical Downlink Control CHannel: The Physical Downlink Control CHannel (PDCCH) carries the Downlink Control Information (DCI) which consists of resource assignments and control information for UEs.

Physical HARQ Indicator CHannel: The role of Physical HARQ Indicator CHannel (PHICH) is to carry the ACKnowledgement (ACK)/Non-ACKnowledgement (NACK) for HARQ, which determines if the eNodeB has received a transmission on Physical Uplink Shared CHannel (PUSCH) correctly or not. For the ACK, the value of the HARQ indicator is 0; while for NACK, it is 1.

The mapping of downlink transport channels to downlink physical channels is depicted in Figure 1.8.

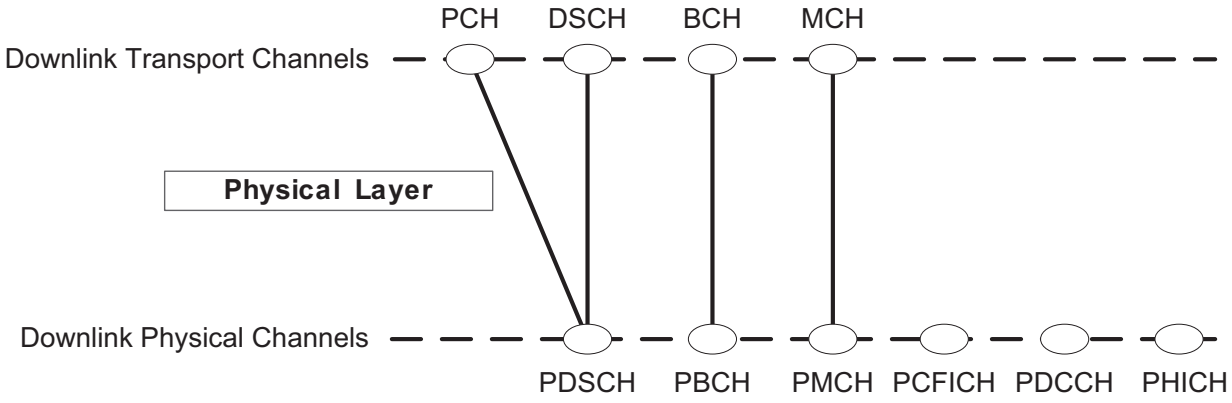

Figure 1.8: Downlink transport and physical channel mapping

1.1.6.2 Uplink Transport and Physical Channels

The uplink transport channels are as follows:

Uplink Shared CHannel: The Uplink Shared CHannel (USCH) supports HARQ, AMC, transmit power variation and dynamic resource allocation. USCH transports uplink user data and control messages.

Random Access CHannel: The Random Access CHannel (RACH) provides initial access to the system when the UE is not synchronized accurately for the uplink, the call setup and the exchange of limited control information.

The uplink physical channels are as follows:

Physical Uplink Shared CHannel: The PUSCH carries data from the USCH.

Physical Uplink Control CHannel: The role of Physical Uplink Control Channel (PUCCH) is to carry the ACK/NACK for HARQ if the data transmitted from the eNodeB is received at the UE correctly (or not). It also carries the scheduling requests and the Channel Quality Indicator (CQI) report.

Physical Random Access CHannel: The initial network access is provided by the Physical Random Access CHannel (PRACH).

The mapping of uplink transport channels to uplink physical channels is depicted in Figure 1.9.

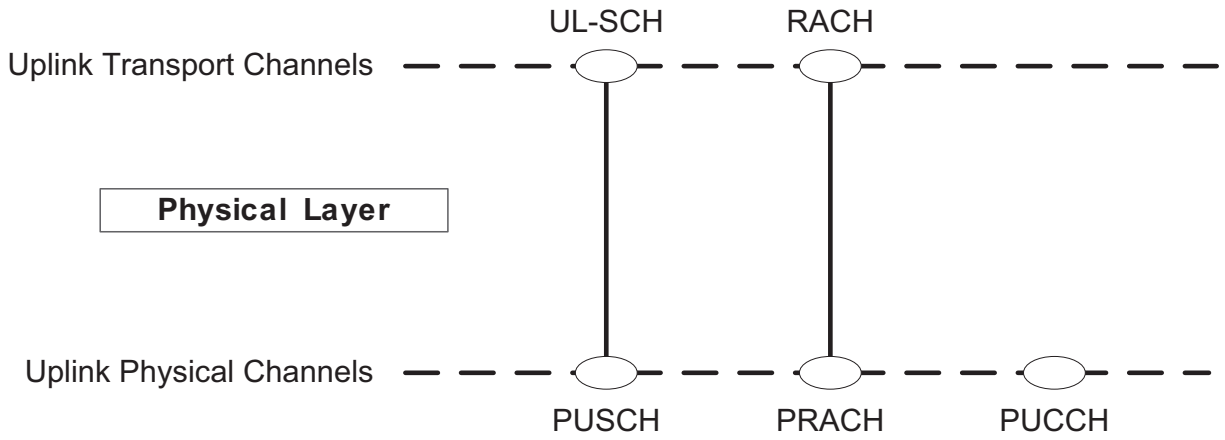

Figure 1.9: Uplink transport and physical channel mapping

1.1.7 Admission Control

Admission Control (AC) is a decision making functionality of the RRM located at the eNodeB. The AC determines whether to accept or reject a new radio bearer request or a handover request. The decision regarding such a request is based on whether efficient radio resource usage is possible if the request is accepted. If the QoS requirements of existing bearers cannot be fulfilled due to admission of a new bearer, or the new bearer cannot be served according to its requirements, the new request has to be rejected. The AC must have QoS awareness in order to make sure the efficient use of radio resources if a new request is made. In a nutshell, AC grants admission to new requests if resources are available and existing users are not expected to suffer from this grant of admission. The AC decisions rely on local cell load information and also information received by the eNodeB from the neighboring cell. QoS provision is ensured by the AC with the help of ARP of the newly requested bearer.

1.1.8 Hybrid Automatic Repeat reQuest

LTE supports both ARQ and HARQ functionalities. In ARQ, data is transmitted after adding redundant bits to it using error detection codes such as the cyclic redundancy check. However, in HARQ, a combination of forward error correction code and error detection code are added to the signal before transmission. The ARQ is located in the RLC layer of E-UTRAN protocol stack, while the HARQ is located in the MAC layer. HARQ monitors the exchange of data packets between the peer entities of the physical layer of the eNodeB and the UEs. In case of erroneous delivery, HARQ makes sure that a physical layer retransmission occurs. A

given number of retransmissions are attempted each time a delivery of a particular packet fails. HARQ makes the system robust and reliable against errors such as inaccurate channel estimation.

The HARQ functionality depends on ACK or NACK messages from the receiver side. Each data packet has to be acknowledged with an ACK or NACK. The ACK corresponds to a successful transmission and a NACK corresponds to a transmission failure. In case of a NACK message, a retransmission is requested. Data is retransmitted using either chase combining or incremental redundancy. In chase combining, an identical copy of the packet is retransmitted which is combined with the incorrectly received copies to increase the Signal to Noise Ratio (SNR) over the information contained. In case of incremental redundancy, additional redundancy is incrementally sent to achieve the decoding of the packet.

1.1.9 Uplink Signaling

The purpose of signaling in the LTE uplink is to provide the channel, buffer and power information to the eNodeB. The eNodeB has to be channel aware to perform efficient resource allocation. The information regarding buffer status of UEs is also significant, so that resources are allocated accordingly. The UE transmit power per PRB information should also be available at the eNodeB so that the Adaptive Transmission Bandwidth (ATB) and Packet Scheduling (PS) functionalities can be facilitated. A brief description of LTE uplink signaling is provided here.

Channel State Information: The uplink Channel State Information (CSI) is quite significant at the eNodeB. To make this information available at the eNodeB, Sounding Reference Signals (SRS) are transmitted from the UEs. The CSI is estimated by measuring the Signal to Interference and Noise Ratio (SINR) of the SRS. This information is utilized to perform AMC and PS. SRS can be transmitted over the whole bandwidth, or a portion of the bandwidth. CSI can also be utilized for Power Control (PC) functionality.

Buffer Status Report: Buffer Status Report (BSR) signals are sent from the UEs towards the eNodeB to facilitate the radio resource allocation at the eNodeB. The general scheduling principle is to allocate resources to UEs if there is any pending data in the data buffers. In downlink, the eNodeB has the knowledge of user data because the data buffers are located in the eNodeB. In uplink, the buffers are located in the UEs and the scheduling is performed by the eNodeB. This information is acquired by the eNodeB using BSRs. The eNodeB should be aware of the amount of data in the buffers of the UEs so that the resources are allocated

accordingly. In order to facilitate QoS based radio resource allocation, separate BSRs should be sent to the eNodeB for each radio bearer.

Power Headroom Report: The Power headroom report is sent by UE to its serving eNodeB in order to facilitate various RRM functionalities, packet scheduling in particular. In the power headroom report, the UE sends the information about the power per resource element or the Power Spectral Density (PSD) to the eNodeB. The eNodeB has to determine how many PRBs can be allocated to a UE so that the maximum power constraint of the UE is respected. The resources are allocated only to a UE capable of utilizing them. This helps in efficient usage of radio resources as the bandwidth which is utilizable by a UE is made available to other UEs.

1.1.10 Power Control

CDMA communication systems are typically characterized by intracell and intercell interference problems. The primary source of interference is the intracell interference between users in a cell because of the time and frequency resource sharing by the users. This interference is reduced by the fast PC scheme. On the other hand, the OFDM based systems are designed in such a way that intracell interference is totally eliminated (at least theoretically). This is the result of the orthogonality of the subcarriers. The aim of PC in the OFDM systems is to overcome the slow channel variations and interference to other cells. Therefore, such a scheme would be a slow PC scheme. In brief, PC in LTE

- provides the users with the required SINR;
- controls the intercell interference;
- minimizes battery power consumption;
- minimizes the dynamic range at eNodeB receiver.

In CDMA systems, users can transmit using the complete bandwidth. In LTE, a user is able to transmit with only a portion of bandwidth. Therefore, the PSD, which is the transmit power per PRB, is adjusted for users and not the total transmit power. Once the PSD is defined for a user, it is kept constant over all the allocated PRBs. Users with bad channel conditions would be penalized by reducing the transmit bandwidth for them and not by reducing the Modulation and Coding Scheme (MCS) level. This would, however, result in only a marginal performance loss [XRG+06].

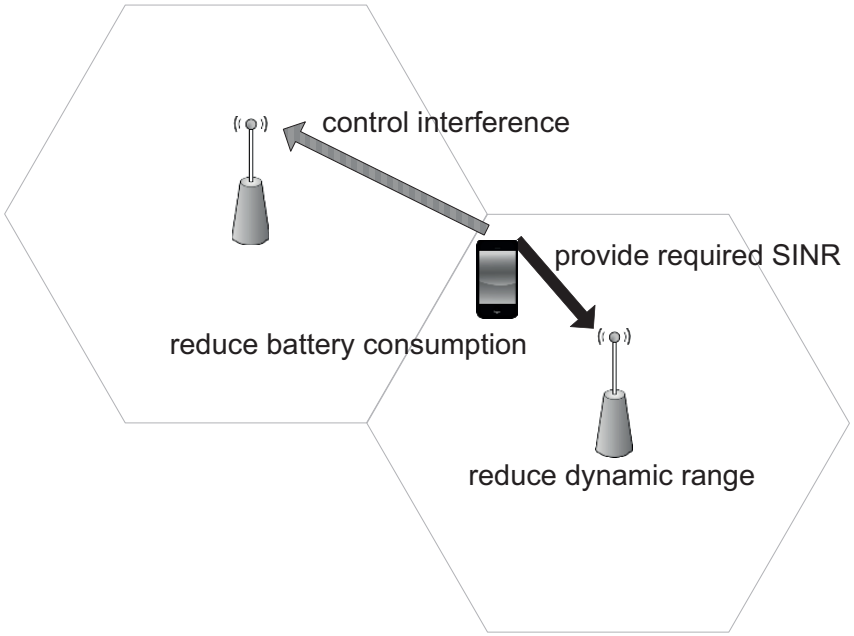

Figure 1.10: Power control advantages

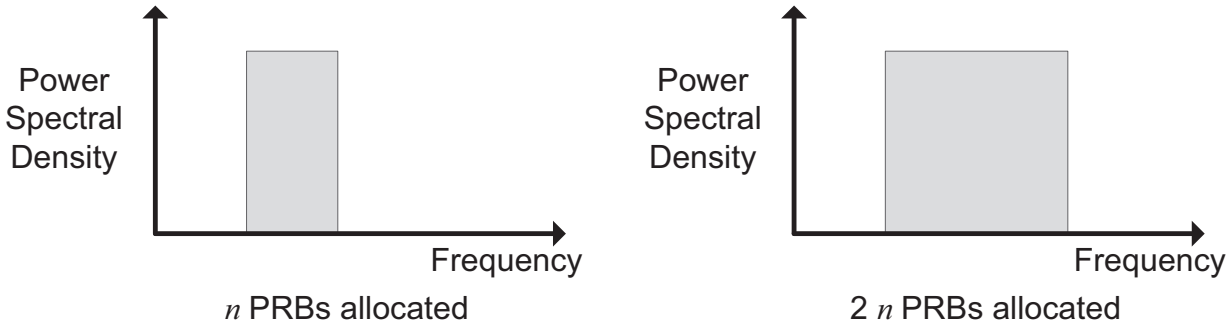

Figure 1.11: Uplink transmit power with different allocations of bandwidth

Uplink PC is performed in the Physical Layer of the UE. PC achieves the goal of providing the required SINR to the UE by modifying the transmit power at the UE; in other words, compensating for user path loss. If the path loss has to be completely compensated, then all the users would be received with the same SINR at the eNodeB and a large eNodeB dynamic range could be avoided. This is known as full path loss compensation. If no PC is employed (i.e. no path loss is compensated), all the users would transmit with the same power and would be received at the eNodeB with varying SINR. In case of complete compensation, a UE with higher path loss would have to transmit with higher power. It would result in generating interference towards neighboring cells, which is a typical cell-edge scenario. To tackle this issue, 3GPP has introduced Fractional Power Control

(FPC) [36.10a]. Under FPC, the users have to compensate for only a fraction of path loss. The cell-edge users with higher path loss would not be required to transmit with very high power and therefore, reduce the interference generated towards other cells.

The FPC formula for PUSCH transmission, adopted by 3GPP, is based on a combination of Open Loop Power Control (OLPC) and Closed Loop Power Control (CLPC) algorithms. The OLPC algorithm is designed to compensate for the slow channel variations. The UE power setting is based on measurements obtained from signals transmitted by the eNodeB, but no feedback is sent to the eNodeB regarding transmit power. To counter the intercell interference and the path loss measurement errors, the CLPC component of the FPC scheme can be used to perform power adjustments based on UE feedback sent to eNodeB. The transmit power according FPC for PUSCH transmission is determined as

$$P = \min(P_{\max}, P_0 + \alpha L + \Delta_{MCS} + f(\Delta_i) + 10\log_{10}N) \tag{1.1}$$

where $P_{\max}$ is the maximum allowed UE transmit power, P_0 is a user specific (or optionally cell specific) parameter expressed in [dBm/PRB], α is the cell specific path loss compensation factor with possible values $\{0, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, L is the downlink path loss measured at the UE using the Received Signal Received Power (RSRP) of the downlink reference signals from the eNodeB, Δ_{MCS} is a user specific (or optionally cell specific) parameter signaled by the upper layers, Δ_i is a user specific closed loop correction value, $f(.)$ is a function for the CLPC command transmitted by eNodeB for correction and N is the number of PRBs allocated to a user. If the maximum allowed transmit power for the UE is ignored, the equation can be analyzed as in (1.2)

$$P = \underbrace{P_0 + \alpha L}_{\text{basic open loop operation point}} + \underbrace{\Delta_{MCS} + f(\Delta_i)}_{\text{dynamic offset}} + \underbrace{10\log_{10}N}_{\text{bandwidth}} \tag{1.2}$$

If the CLPC component is ignored, then the basic OLPC scheme can be formulated as

$$P = P_0 + \alpha L + 10\log_{10}N \tag{1.3}$$

The parameters like P_0 and α have to be carefully tuned to accomplish throughput and outage gains. It is shown in [ZMS+08] that the optimum value for $\alpha = 1$ is $P_0 = -106$dBm and optimum value for $\alpha = 0.6$ is $P_0 = -58$dBm.

In order to use the PC functionality, the Interference based Power Control (IPC) scheme proposed in [BQC+08] is utilized. The user transmit power in this algorithm is adjusted not only by considering the path loss to the serving cell but also the interference generated towards the other cells in the network. This interference is the path gain (reciprocal of path loss, $G = -L$) of the user towards other cells. The authors of [BQC+08] have evaluated the scheme in terms of the outage user throughput and the average cell throughput. In comparison to the FPC scheme, IPC has been shown to provide substantial outage and average cell throughput gains with careful parameter settings. IPC has been formulated as

$$P = P_0 + \alpha L + \beta L_{other} + 10\log_{10} N \tag{1.4}$$

where β is the function of interference towards other cells (similar to α) and L_{other} is the logarithmic sum of path losses towards other cells $(L_1 + L_2 + L_3 + ...)$. It has been shown that desired results are achieved if parameters are tuned such that $\alpha + \beta = 1$. More specifically, the combinations $\alpha = 0.6$, $\beta = 0.4$ and $\alpha = 0.1$, $\beta = 0.9$ provide best results [BQC+08].

PC is only employed in the LTE uplink; there is no PC for the LTE downlink, where the transmission power strategy is to transmit with a constant output power, i.e. maximum eNodeB power.

1.1.11 Packet Scheduling

Scheduling of radio resources in LTE is performed by the MAC scheduler of the eNodeB. The scheduler is designed to distribute radio resources among the radio bearers in a cell. The downlink radio resources are allocated to a UE if downlink data is present in the buffer of the eNodeB for that particular UE, enough resources are available and the QoS requirements of other users can be fulfilled. Resources are allocated to a UE for uplink if data is present in the UE buffer. This information is sent by the UE to the eNodeB using the BSR (Buffer Status Report). Transmission of the BSR from the UE to eNodeB is very important for the scheduler because the scheduling has to be performed at the eNodeB and data buffer is located in the UE. Therefore, by sending the BSR, the UE requests the scheduler to provide resources. the BSR also contains information about the type of traffic in the UE buffers. The eNodeB has to allocate resources by guaranteeing the QoS requirements of the radio bearers involved in the downlink and the uplink.

In LTE, all kinds of traffic is sent over the LTE network in the form of packets. The term ‘packet scheduler’ is used for schedulers that facilitate the packet switched network with the resource allocation functionalities. In a packet switched

network, requests for the downlink and the uplink scheduling are not served on first-come first-served basis. Certain priority metrics are used by the packet scheduler to decide which requests are to be served with higher priority than the others. Some commonly used algorithms for priority metric generation are

- Blind Equal Throughput (BET)
- Maximum Throughput (MT)
- Proportionally Fair (PF)

Besides these simple scheduling algorithms, advanced and complex algorithms can also be employed in LTE. Algorithms which consider the instantaneous channel conditions of a user, QoS requirements or UE buffer status can be effectively used to increase system throughput and spectral efficiency. Channel fading information can also be exploited positively by the packet scheduler [Pok07]. If the scheduler is channel aware, it can utilize time and frequency diversities in order to increase the spectral efficiency. Time diversity can be utilized by postponing resource allocation to users in deep fade until recovery of the link quality. Frequency diversity can also be utilized by the scheduler if a user is allocated its PRBs with good signal quality and allocation of highly attenuated PRBs is avoided. Previously, fading was considered a serious limitation in performance of a wireless communication system.

Dynamic Packet Scheduling: One of the significant features of LTE scheduling is the ability to schedule packets dynamically among the downlink or uplink bearers. This special feature of LTE in combination with Link Adaptation (LA) can enable the system to achieve a high spectral efficiency and fulfill the QoS requirements of the users in a cell. The dynamic packet scheduler located in the MAC layer of the eNodeB distributes the PRBs among the users by choosing an appropriate MCS. These tasks are performed by the scheduler in every Transmission Time Interval (TTI) or 1ms and therefore, termed as fast scheduler [Cal09]. The length of a TTI is 1ms.

The scheduling information is broadcasted to all the UEs through the PDCCH. UEs have to monitor the PDCCH to know when to transmit and receive. In downlink, the scheduling decisions are made on a per bearer basis. The bearers of a single user can be treated independently, based on their QoS requirements. In uplink, however, the scheduling decisions are made on per user basis. A user with high priority is scheduled before the low priority users. The priority of the users is defined with the help of the algorithms used by the scheduler.

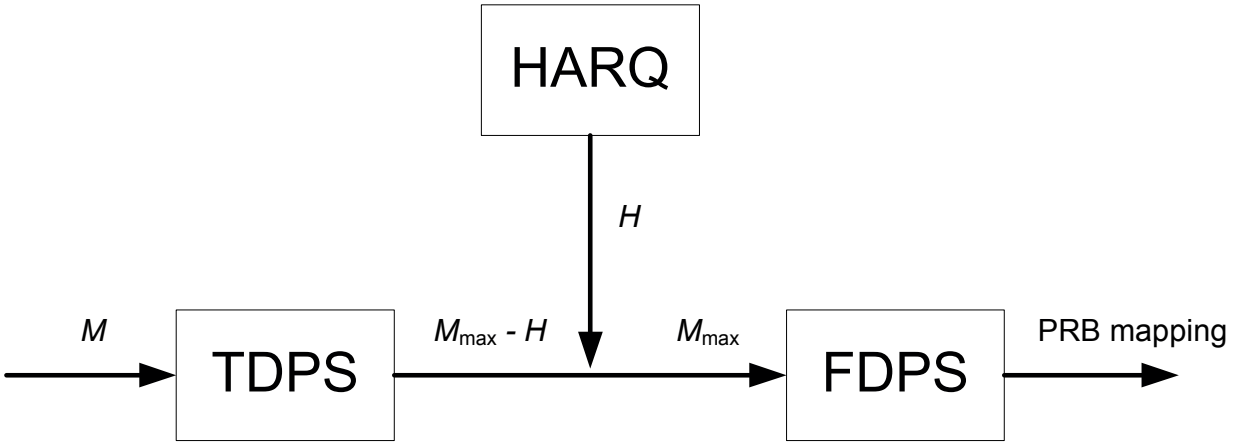

Figure 1.12: Simple scheduling block diagram

Time Domain Packet Scheduling: Time Domain Packet Scheduling (TDPS) is the 1st phase of packet scheduling. This phase has the task to choose the users which are to be scheduled within the TTI. The users are selected on the basis of TDPS metrics. The TDPS metric value generated via algorithms is based on certain criteria, e.g. maximizing throughput, serving QoS requirements etc. The users selected by TDPS are forwarded for bandwidth allocation in the frequency domain.

Frequency Domain Packet Scheduling: The 2nd phase of packet scheduling is the Frequency Domain Packet Scheduling (FDPS). In FDPS, only the users selected by TDPS are considered for scheduling and resource allocation. The computational complexity of FDPS is reduced by TDPS because the number of users allowed to enter FDPS is limited by TDPS. The computational complexity of FDPS increases as the number of users is increased. This implies that all the users having uplink data may not be allowed by the TDPS to enter FDPS. The task of FDPS is to map the selected users to PRBs. PRB mapping is based on FDPS metrics. Figure 1.12 illustrates the functionality of scheduler. If M is the number of users requesting resources, $M_{\max}$ is the maximum number of users allowed to enter FDPS and H is the number of users with pending HARQ retransmissions, then the number of users going from TDPS to FDPS would be $M_{\max} - H$.

Downlink and Uplink Scheduling Comparison: As explained, the downlink scheduler allocates radio resources to users based on the data buffered at the eNodeB. Since OFDMA is the radio interface for downlink in LTE, the resource allocation is performed with great flexibility. The downlink scheduling can be divided into

phases like QoS based bearer classification, time domain packet scheduling and frequency domain packet scheduling.

In QoS based bearer classification, the data buffered in the eNodeB for downlink are classified according to QoS classes. All the bearers of all the users are prioritized according to the TDPS algorithm used by the eNodeB. Similarly, the PRBs of all the users are prioritized according to the FDPS algorithm. TDPS is performed on per bearer basis and the bearers of a user are treated independently, with respect to their QoS requirements. In TDPS, the bearers with high priority metrics are chosen for frequency domain scheduling. TDPS determines which bearers are to be facilitated by FDPS. In FDPS, the selected users are mapped to PRBs and resources are allocated accordingly.

In uplink scheduling, SC-FDMA is the radio interface scheme chosen for LTE. SC-FDMA also utilizes the capability of OFDM to allocate the radio resources flexibly. Uplink scheduling can be divided into several phases such as user classification, TDPS, FDPS and user bearer scheduling.

In the user classification phase, users are classified and priority metrics are assigned according to the TDPS algorithm. The metrics can be based on QoS requirements of user, channel state or buffer status. Similarly, the PRBs of all the users are also prioritized according to the FDPS algorithm. TDPS is performed on per user basis in uplink based on the BSR sent by the UE to the eNodeB. In TDPS, the users with high TDPS metric values are chosen for FDPS. FDPS in uplink is more complicated compared to the downlink because of the constraint of assigning contiguous PRBs to a user due to the SC-FDMA radio interface. FDPS in uplink also has to deal with the power constraints of a user for efficient use of battery in the UE. This constraint means that user may not be able to obtain greater bandwidth for compensation of bad channel conditions. The uplink scheduler should be designed taking all these features and constraints into consideration. This makes the uplink scheduler designing a challenging task.

1.1.12 Link Adaptation

LA is an important LTE functionality used to adjust the transmission rate of data according to user channel conditions. The user channel conditions are made available at the eNodeB by sending it the CQI report from the UE. In uplink, the channel conditions are provided to the eNodeB by transmitting the SRS to it from the UE. Based on this information, the eNodeB performs LA by adjusting the modulation scheme, coding scheme, transmission bandwidth and transmit power for each UE. Some of these adjustments are explained here in brief (excluding PC, which has already been explained in detail).

Adaptive Modulation and Coding: The AMC functionality is used to enhance the spectral efficiency by exploiting the user channel conditions. The user SINR can be determined by using the Effective Exponential SINR Mapping (EESM) method in downlink and Average Value Interface (AVI) method in uplink to obtain a system level SINR value for each user. This SINR value is mapped on to an MCS using mapping tables. The MCS is used to determine the TBS for each user. The LTE modulation schemes are QPSK, 16-QAM and 64-QAM.

Adaptive Transmission Bandwidth: The ATB functionality denotes the bandwidth scalability feature of LTE which enables the allocation of varying bandwidths to users. Some traffic types can be served with small bandwidth, while others require greater bandwidth. Varying cell load also demands that the bandwidth is allocated flexibly depending on the number of users. Moreover, the adversity of user channel conditions may limit its bandwidth. In all these cases, it would be required that the bandwidth is adjusted accordingly. The reduction of bandwidth for one user can increase the throughput of other users, hence improving the overall system performance.

1.2 LTE-Advanced

As an enhancement to LTE, the LTE-Advanced (LTE-A) standard was presented as a candidate fourth generation (4G) system to the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) in 2009, and documented by the 3GPP in Release 10 and 11. It is expected to keep backward compatibility with LTE, which means that the LTE-A network can be deployed in the frequency bands occupied by the LTE system and would thus be able to utilize most of the LTE technologies. Meanwhile, the LTE terminals can also work in the LTE-A system.

In order to achieve a better performance such as a higher data rate and a better throughput, LTE-A has introduced several new features in addition to LTE features. These features include the Carrier Aggregation (CA), the enhanced multiple-input and multiple-output (MIMO), and the Relay Node (RN) etc.

1.2.1 Air Interface

As already explained, OFDMA is the LTE downlink air interface scheme, providing wideband transmission while staying robust to frequency selectivity of radio channels. The uplink uses SC-FDMA as the transmission scheme due to the low

PAPR as compared to the OFDMA signal. However, the SC-FDMA used in LTE has a constraint that only allows adjacent radio resource allocation in the frequency domain.

In LTE-A, OFDMA is also used for downlink transmission. For LTE-A uplink, due to several types of CA, the uplink air interface turns out to be the aggregation of multiple SC-FDMA bands, which is denoted as Aggregated Discrete Fourier Transform-Spread OFDMA (n*DFTS-OFDMA). This aggregation results in the air interface being able to support both contiguous and non-contiguous resource allocation in uplink to user terminals. DFTS-OFDMA, thus, overcomes the disadvantages of the OFDMA (high PARP) and the SC-FDMA (contiguity) in LTE.

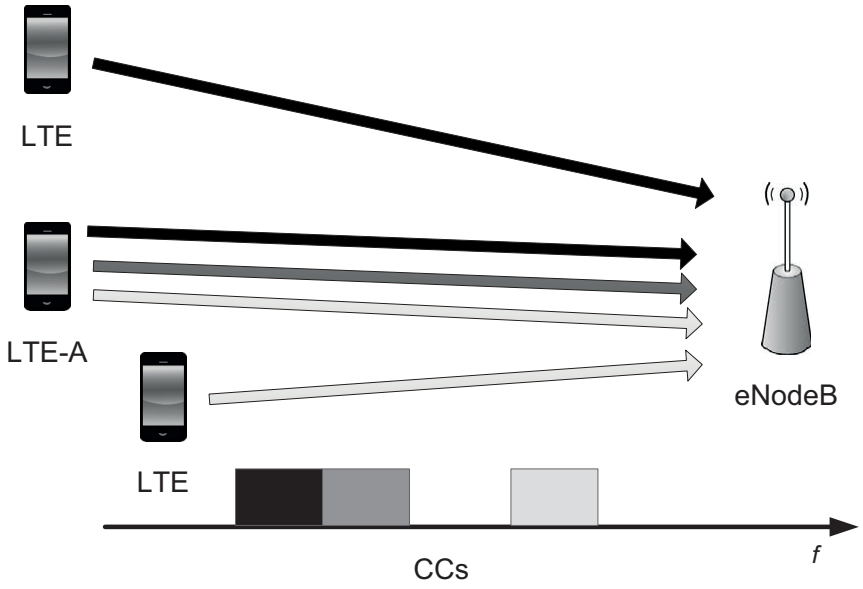

Figure 1.13: Carrier aggregation

1.2.2 Carrier Aggregation

To achieve higher data rates, a wider bandwidth is required for data transmission. According to the 3GPP specifications [36.10b], LTE-A is capable of aggregating two or more Component Carriers (CCs) to obtain a wider transmission bandwidth. This is defined as CA. The current 3GPP standards allow for up to five CCs to be aggregated. LTE terminals can only receive or transmit on a single CC. Whereas LTE-A terminals can receive or transmit data on one or multiple CCs simultaneously Figure 1.13. CA can be classified into three types: the intra-band contiguous, the intra-band non-contiguous and the inter-band non-contiguous Figure 1.14.

In LTE, a UE can only have one RRC connection with the eNodeB. However in LTE-A, several serving cells can be formed due to CA, with each cell corre-

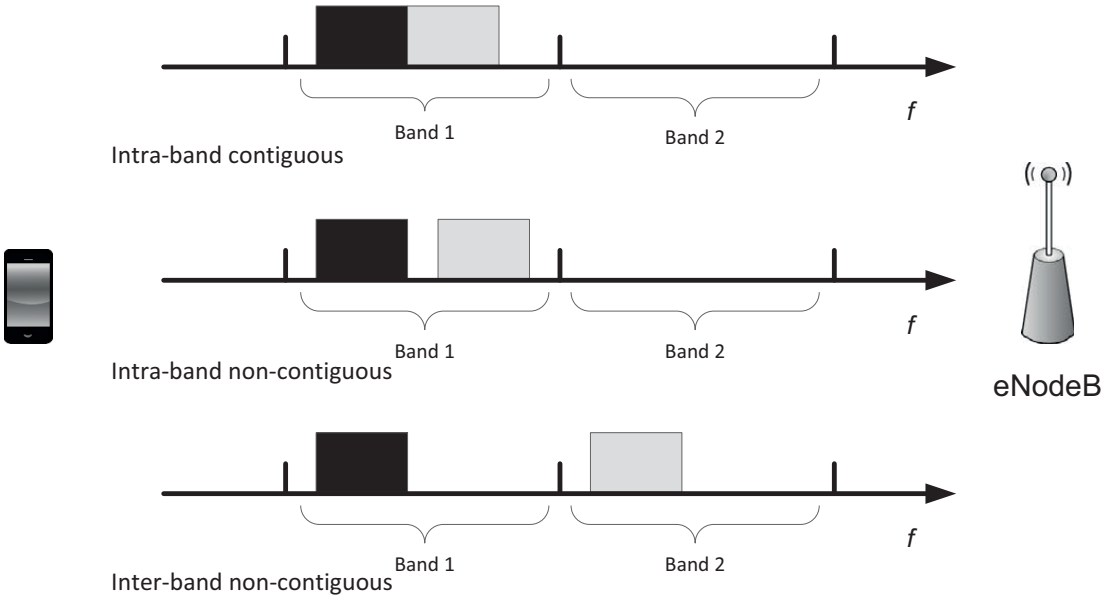

Figure 1.14: Carrier aggregation types

sponding to a CC. An RRC connection is initiated with a single CC called Primary Component Carrier (PCC), using the same RRC establishment procedure as specified for LTE. Afterwards, additional CCs can be configured from the eNodeB. An additional CC is called Secondary Component Carrier (SCC). The SCCs can be added and removed as required, whereas the PCC is only changed at handover. For UEs using the same set of CCs, different PCCs are possible (Figure 1.15). The PCC serves the primary serving cell and the SCCs serves the secondary serving cells.

The Component Carrier Selection (CCS) scheme allocates one or several CCs to the UEs according to certain criteria. In downlink, the allocation of multiple CCs results in a higher throughput and large transmission bandwidth as there is no power constraint. However, this might not be the case for the uplink, especially for the UEs who are power limited at the cell edge. Therefore, efficient methods for CC selection are required. The uplink packet scheduling scheme should be designed in such a way that the scheduler is aware of the CC selection.

1.2.3 Coordinated MultiPoint

Future mobile networks are expected to serve a large number of mobile terminals simultaneously. Traditionally, each mobile terminal is assigned to one eNodeB, depending on signal strength, terminal distance to the eNodeB and so on. Ideally, there should be no interference within the same cell as the OFDM scheme makes

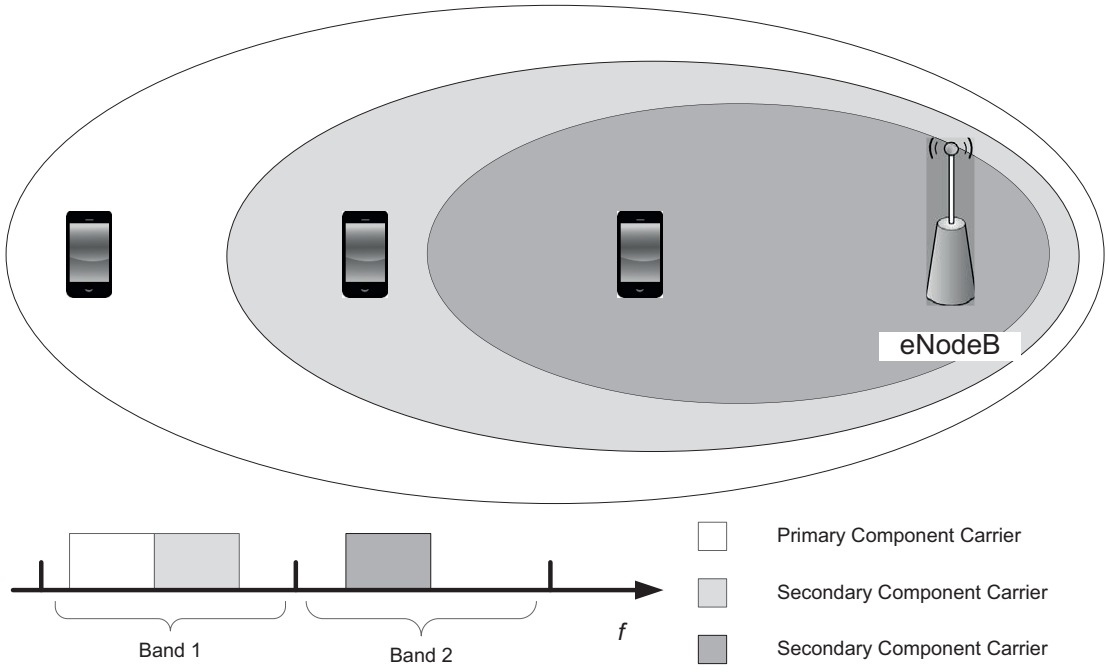

Figure 1.15: Primary and secondary cells

sure that the signals are orthogonal to one another. However, signals from the other eNodeBs can interfere with the mobile terminals. In addition, the UEs in the serving eNodeB also causes interference to the UEs in the other eNodeBs. This interference can be reduced by using Coordinated MultiPoint (CoMP) transmission/reception [36.10b]. In CoMP, multiple geographically distributed antennas interact with each other to improve the terminal performance. The specifications define CoMP techniques for both downlink and uplink.

1.2.4 Relay Nodes

The Relay Node (RN) feature is introduced by 3GPP [36.12] to improve the system performance of LTE-A. The RNs are low power nodes used to improve network performance by increasing the network coverage and cell edge throughput. An RN is connected to a Donor eNodeB (DeNB) via the Un air interface in the E-UTRAN. The DeNB not only serves its own UEs, but also shares radio resources with the RNs. Figure 1.16 depicts the functionality of the RN. The details of various RN functionalities are discussed in chapter 4.

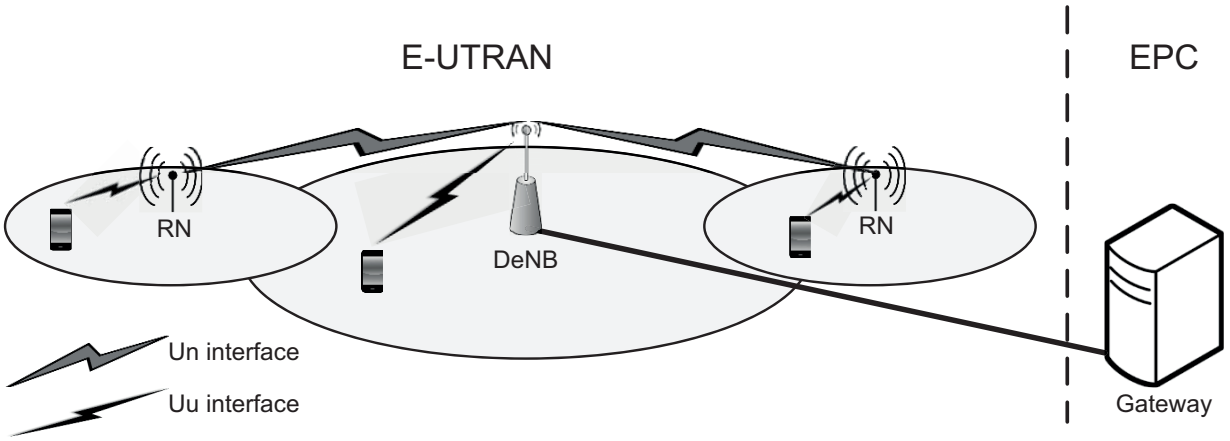

Figure 1.16: Relay node for coverage and cell edge throughput improvement

1.3 M2M Communication

Machine-to-Machine (M2M) communication is one of the latest research areas and has received tremendous endorsement globally. The crux of M2M communication is enabling the devices i.e. machines, including computers, sensors and actuators to interconnect with each other without human intervention [WPSM04]. The prices and sizes of such devices have seen a significant reduction lately, leading the manufacturers to offer new applications and services. It has been forecasted that M2M based mobile data traffic is going to see exponential growth in the near future [Cis14]. M2M communication is described in detail in Chapter 3.

1.4 Problem Statement

Prevalent M2M communications are reliant on available mobile communication networks like GSM. The contemporary networks are capable of fulfilling the requirements of existing M2M services satisfactorily for the time being due to the low cost deployment with roaming facilities of these networks. In spite of that, the anticipated growth of M2M traffic implies that the legacy mobile systems would soon be considered insufficient. It is important to note that a main part of M2M traffic differs from typical human based video or web browsing type of traffic in the sense that M2M traffic originates from a huge number of devices in the form of small data messages. However, the human based regular traffic emerges from a limited number of devices with frequent data transmission or large sized messages.

LTE-A is considered to be a candidate for serving the massive future M2M traffic. M2M applications are generally designed with narrowband requirements.

However, the development of 3GPP standards is mainly for serving broadband data traffic. These standards are not able to achieve spectrum and cost efficiency with the current architecture and design for narrowband M2M applications. Hence, incorporating the narrowband M2M data traffic with small packet sizes and higher number of devices require novel modifications in the design of the 3GPP based networks to avoid substantial degradation to the overall network performance.

In order to achieve spectral efficiency of the LTE-A network air interface, the scarce radio resources should be utilized efficiently. A PRB is the smallest resource allocation unit for a device. In regular human based LTE-A traffic, a PRB can be efficiently utilized by transmitting one or several packets of user data with a single PRB. On the other hand, a PRB cannot be fully utilized for narrowband M2M applications because the devices transmit small sized messages with a transmission gap of one or several seconds. Each device has to use at least one PRB for transmission and the PRB cannot be shared by multiple devices. So the radio resources, despite availability are not utilized efficiently.

In this thesis, a proposal is presented for the solution to the problem of utilizing the air interface resources efficiently. Instead of allowing M2M devices direct access of the eNodeB, an intermediate node can be used to aggregate the data of M2M devices and thus, multiplex the aggregated data before transmission over the air interface. RNs are chosen as the aggregation and multiplexing node for M2M in mobile networks in this thesis due to the low power and low implementation costs. RNs can be deployed without any need of additional infrastructure for connecting with the core network, because RNs are connected wirelessly to the DeNB. Although, RNs are primarily designed to enhance cell edge coverage and throughput for regular LTE-A traffic, this thesis explores a method for aggregating and multiplexing M2M traffic at the RN before sending the data to the DeNB.

1.5 State-of-the-Art

The literature survey of RNs indicates that most of the research work focuses on the implementation aspects and improving system performance through relaying. In [ID08], the performance of a relaying system is evaluated by deploying RNs in real environments with propagation models of an urban location in central London. The evaluation is achieved without employing traditional metrics of artificial environments and models. The authors of [RRRH09] investigate the uplink performance of LTE-A with RN deployment and propose a resource allocation scheme according to relaying requirements. In [VWHS09], field trial measurements using a test-bed for indoor relaying are performed with full frequency reuse. The field

trial measurements are extended to outdoor relaying in [WVH+09].

Performance evaluation of relaying in LTE-A has also been performed by several researchers. In [HjNT+10] improvements in relaying protocols and frame structure are suggested to enhance the overall system performance. In [TFP+10], a user data multiplexing scheme for RNs is introduced for reducing the multiplexing overhead. In [BRRH10], the influence of site planning on the performance of relay networks is investigated and the system performance improvement with interference mitigation is achieved with Power Control (PC) schemes at the DeNBs as well as the RNs. Outband relaying operations (explained in chapter 4) are studied in [KYMP12] for power and bandwidth optimization. In [BL12], the authors investigate the reduction in relaying latency by enhancing the available relaying schemes. The QoS-aware scheduling issues for RNs with inband operations is reviewed in [dMBS13]. The role of repeater type RNs for M2M communication is investigated in [ELA11].

Literary work available on packet multiplexing is mostly related to 3G or previous networks. The performance of a statistical multiplexer for determining the delay of voice packets is investigated in [HHL86] by a Markov modulated Poisson process. A packet encapsulation and aggregation process with Poisson arrivals and phase-type service time distribution is used to analyze the end-to-end packet delay performance in [HGSO06]. An approximation method for determining the probability distribution function of the packet waiting time for multiplexed flows is proposed in [MM10].

The Asynchronous Transfer Mode (ATM) is a cell transfer technology and supports various traffic classes such as voice and video. The ATM Adaptation Layer 2 (AAL2) is a multiplexing technology for 3G networks. It is capable of packing small packets into ATM cells. In [HTV03], a stochastic model for the AAL2 layer of the transport network of UMTS Radio Access Network is presented. In [HV06], a queueing model is proposed for the performance analysis of the AAL2 multiplexer in terms of waiting time distribution and multiplexing gain based on a batch Markovian arrival process with timer. This work is further extended in [HT05] for the model of the multiplexer queue. In [CTZC02], an analytical model is presented for the performance of the AAL2 multiplexer to evaluate the mean queue length, mean sojourn time and the delay violation probability. In [JSC11], the AAL2 multiplexer is analyzed by modelling the departure process of the multiplexer as an r-stage Coxian process.

1.6 Contributions of this Thesis

The main motivation behind this thesis is to design a framework for M2M communication over LTE-A network in an efficient manner. In order to achieve this goal, an RN based approach for data aggregation and multiplexing is presented for M2M applications with narrowband requirements.

However, before the design of the proposed framework is described, some of the features and functionalities of the 3GPP networks related to resource management are elaborated in chapter 2 which are vital for this thesis. The LTE uplink radio resource allocation scheme (designed as Master Thesis [Mar11]) is reviewed and the scheduler performance is evaluated with the help of simulation results. This scheme is designed to fulfill the QoS demands of users. The scheduler also considers the channel conditions, power control constraint, fairness and PRB contiguity constraint due to the SC-FDMA air interface of LTE. The work on uplink resource allocation is extended to LTE-A uplink scheduling by incorporating the Carrier Aggregation (CA) functionality into the system design.

Chapter 3 gives an overview of M2M communications. The architecture for M2M communication with its domains is explained and the standardization bodies of M2M communication are also mentioned. An insight of the M2M application areas including logistics and e-healthcare is provides as well. The M2M traffic characteristics and future trends are discussed in the context of challenges that future mobile networks are expected to face. The impact of M2M traffic on the performance of LTE regular traffic is also evaluated.

Chapter 4 provides an outline of features and functionalities of the RN. Various classes of relaying systems based on mobility, technology and air interface usage are discussed. An account of different solutions for serving the M2M application with LTE-A network is presented with pros and cons. The RN based scheme is chosen as a solution for the M2M issues in LTE-A networks. The proposed framework is implemented using the OPNET simulation environment and the implementation details are also provided. Simulation results illustrate the advantages of the proposed framework. The results are validated with the help of an analytical model in chapter 5.

2 Broadband Radio Resource Management

Radio Resource Management (RRM) is the set of algorithms and techniques used to control the usage of the radio interface in LTE [Cal09]. The purpose of RRM is to achieve the efficient use of scarce radio resources and at the same time, provide QoS to the users and seamless mobility experience. These functionalities are hidden from the users but require complex network equipments and UEs [STB09]. Network operators have to deal with complexity and performance trade-off during design phase. Downlink and uplink RRM functionalities have the same goals to achieve but are treated separately because of the difference in their features and characteristics. In chapter 1, an overview of the several LTE and LTE-A RRM functionalities was provided. In this chapter, the work related to the RRM functionalities covered in this thesis is explained.

2.1 LTE Uplink Scheduling

In this section, the LTE uplink scheduling schemes designed during this thesis are discussed. In general, separate algorithms are used for time and frequency domain packet scheduling. However, before going into the scheduler design details, an overview of the channel models used is provided here.

2.1.1 Channel Models

The channel conditions of a UE play a significant role in scheduling decisions. The channel conditions are emulated by channel models for path loss, slow fading and fast fading. The channel models used in this thesis are explained here.

2.1.1.1 Path Loss

Path loss is the power attenuation of the transmit signal in the propagation environment as the signal travels from the transmitter to the receiver. Path loss is a vital factor in channel modelling that cannot be avoided in any propagation environment. The channel model should take into account the power lost due to path loss. Path loss occurs even in a free space propagation environment.

The vehicular test environment model in (2.1) models path loss at a carrier frequency of 2GHz [10198]. This model has been widely used in literature for LTE system level simulations ([WJR12],[Bal10],[KO13] etc.), and also utilized as a model of path loss for uplink.

$$L = 128.1 + 37.6\log_{10} R_{km} \tag{2.1}$$

where L is the path loss in dB and R_{km} is the distance between transmitter and receiver in kilometers.

2.1.1.2 Slow Fading

Slow fading is a long-term amplitude fluctuation of the transmitted signal. Slow fading is primarily caused by shadowing and, therefore, frequently called shadow fading. Shadowing is the obstruction that a signal encounters due to objects blocking its path. The log-normal shadowing model in (2.2) is a common method of modelling slow fading caused by shadowing [Gol05].

$$f(S) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{dB}}} \exp\left[-\frac{(S-\mu_{\mathrm{dB}})^2}{2\sigma_{\mathrm{dB}}^2}\right] \tag{2.2}$$

where $f(S)$ is the probability distribution function of shadowing S in dB (S is the difference between transmit and receive power in dB scale and is random having normal distribution), μ_{dB} is the mean of S in dB and σ_{dB} is the standard deviation of S in dB.

Field tests have shown that log-normal shadowing can generate near-accurate shadow fading effects for the outdoor and indoor reception of signal [MEJA91]. In the frequency domain, shadowing is highly correlated even if carriers are 900 MHz apart from each other. In OFDM, the subcarriers are quite close to each other. Therefore, all the subcarriers encounter approximately the same shadowing effects.

Shadowing measurements at two points within a cell that are at a certain distance apart from each other are presumed to have correlation among them. When modelling shadow fading, this correlation over distance is also modelled. Gudmundson [Gud91] suggested a model which depicts correlation between two points separated by a distance δ. This model can be expressed mathematically as in (2.3).

$$A(\delta) = {\sigma_{\mathrm{dB}}}^2 e^{\delta/d_c} \tag{2.3}$$

where $A(\delta)$ is the covariance of shadowing between two points at distance δ apart and d_c is the decorrelation distance, i.e. the distance at which the covariance between the two point is ${\sigma_{\text{dB}}}^2 e^{-1}$ (where ${\sigma_{\text{dB}}}^2$ is the maximum value that $A(\delta)$ can have). In [10198], ETSI has described that d_c depends on the environment, and suggested the d_c value of 20m for the vehicular test environment and 5m for the indoor environment.

This shadowing correlation over distance model has to be integrated into the log-normal shadowing model as in [CYG95]. For such an integrated model, a moving mobile user is considered to be positioned at an initial point x_0 where the shadow fading value is to be randomly generated using the log-normal distribution in (2.2). The dB value of the shadowing at distance 0m from point x_0 is equal to S(0). To determine the shadowing at points which are at a distance of δ, 2δ, 3δ... away from x_0 and δ is equal to 5 meters, then according to [CYG95]

$$S(n.\delta) = e^{\delta/d_c} S((n-1)\delta) + V_i \tag{2.4}$$

where V_i are independent identically distributed normal random variables with mean 0dB and variance ${\sigma^2}_{\text{dB}} [1 - \exp(-2\delta/d_c)]$, while $n = 1, 2, 3, ...$

2.1.1.3 Fast Fading

Fast fading is a short-term amplitude fluctuation of the transmitted signal. Fast fading causes the signal amplitude to fluctuate for a shorter period of time. The scale of fading effect is smaller as compared to path loss and slow fading. Path loss, slow fading and fast fading are superimposed to obtain the overall signal fading gain. Fast fading causes frequency and time selectivity of a signal. Multipath propagation characterizes the reception of a radio signal at the receiver from more than one path. The impulse response is usually a chain of pulses. Additionally, the real channels are time variant with transmitter, scatterers and receiver in motion. Frequency selectivity occurs due to multipath propagation of the signal and time selectivity occurs due to the motion of transmitter, scatterers and receiver.

The Jakes-like method of complex gain generation [Cav00] assesses the gain of the channel by taking into account the Doppler spread for time selectivity and the delay spread for frequency selectivity. The time selectivity, which is caused by Doppler spread, is modelled according to Clarke's model [TV05]. Since the frequency selectivity depends on the carrier frequency, therefore, the channel gain is modelled as a function of both time and carrier frequency. The fast fading chan-

nel gain can be illustrated by a 3-dimensional plot with its axes representing time, frequency and channel gain as in Figure 2.1.

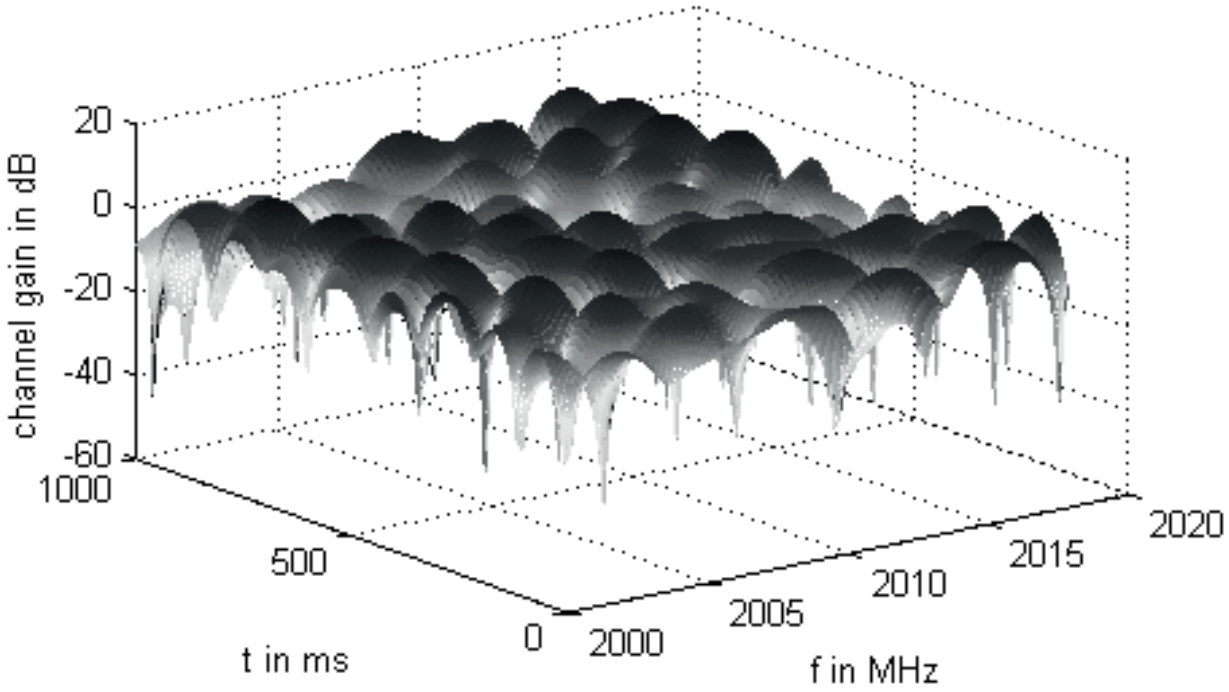

Figure 2.1: Fast fading channel gain versus time and frequency

2.1.2 Scheduler Overview

Most of the available research work on uplink schedulers lacks the consideration of characteristic features of the LTE uplink. For example, several researchers present uplink schedulers without the constraint of allocating resources contiguously to users. Others tackle this constraint problem by allocating equal bandwidth to all the available users. In most of the proposals, QoS requirements of the users are not considered. Others consider the user QoS but assume that each user can have only one class of traffic. With all these issues, the Bandwidth and QoS Aware (BQA) uplink scheduler is proposed in [Mar11] which considers the user QoS requirements, maximum allowed transmit bandwidth, channel conditions and fairness among users. Multi-bearer users are supported and efficiently handled according to QoS requirements. The UE power efficiency improvement and intercell interference reduction is achieved by scheduling according to PC commands for UE transmit power adjustments. The bandwidth allocation is performed using a tree-based algorithm which considers the contiguity constraint for PRB allocation. The transmit PSD of a UE determined by PC, and the buffer size of the UE are

also considered for resource allocation. The bearers within the UEs are scheduled according to their QoS requirements.

The scheduling is performed separately for each cell. Uplink and downlink scheduling is also performed independently of each other in each cell. The uplink scheduler initiates scheduling by identifying the active users in a cell. The users are classified and prioritized according to their channel conditions, QoS requirements and fairness by TDPS. The top priority users are allowed to enter the next phase i.e. the FDPS. In FDPS, the resources are allocated to the high priority users in a proportionally fair manner using the tree based algorithm with QoS consideration. The allocated resources are utilized by the radio bearers of the users according to their QoS requirements.

2.1.3 Time Domain Packet Scheduler

The TDPS is the initial phase of the uplink packet scheduling of the users. The presence of data packets in the buffer of the UE is reported to the eNodeB using BSRs. The BSRs include information about the status of the bearer buffers within a UE. A list of users with pending uplink data is generated. The channel conditions of the active users are also acquired using CSI which is carried by SRS. In this thesis, it is assumed that the SRSs are received at the eNodeB at each TTI and for all the PRBs of active users. Moreover, it is also assumed that eNodeB is aware of the PSD of each active UE with the help of power headroom reports.

2.1.3.1 Time Domain Metric Algorithms

Time domain algorithms in uplink are designed to select high priority users for frequency domain scheduling. Therefore, a user is selected (or rejected) based on the TDPS metric value. The metric value of a user can be derived depending on several criteria. Some commonly used algorithms and the algorithm proposed in this thesis are explained here.

Blind Equal Throughput (BET): In time domain, the BET algorithm attempts to provide equal throughput to all the users, regardless of the user channel conditions [MPKM08]. This algorithm gives no special priority to the users close to the eNodeB. It is an extremely fair algorithm and cell edge users also benefit equally. This fairness has to be delivered at the expense of low cell throughput. If $\Lambda_i(t)$ is the TDPS metric value of user i, this metric value can be given as

$$\Lambda_i(t) = \frac{1}{R_{\text{avg},i}(t)} \tag{2.5}$$

where $R_{\text{avg},i}(t)$ is the average throughput of user i at time t over a time window determined using the Exponential Moving Average (EMA) as in (2.6).

$$R_{\text{avg},i}(t) = (1 - \frac{1}{T_w})R_{\text{avg},i}(t-1) + \frac{1}{T_w}R_{\text{ach},i}(t) \tag{2.6}$$

where T_w is the EMA time window length in TTIs and $R_{\text{ach},i}(t)$ is the actual bit rate achieved by user i in the previous TTI. $R_{\text{ach},i}(t)$ is zero if i is not scheduled in the previous TTI.

Maximum Throughput (MT): The goal of the time domain MT is to prioritize radio resources to the users with good channel conditions. Users close to the eNodeB benefit from this algorithm. Cell edge users get resources on a low priority basic. This algorithm provides maximum cell throughput but at the expense of fairness.

$$\Lambda_i(t) = R_{\text{inst},i}(t, n_i) \tag{2.7}$$

where $R_{\text{inst},i}(t, n_i)$ is the instantaneously achievable wideband throughput of user i at time t with maximum number of obtainable PRBs n_i (which is set according to the PC scheme).

Proportionally Fair (PF): The BET and MT algorithms can be considered to be on two extremes, i.e. absolute fairness and maximum throughput. An optimum algorithm which respects both these considerations and has been extensively used in literature is the PF algorithm. This algorithm attempts to provide certain degree of fairness as well as maximize the cell throughput.

$$\Lambda_i(t) = \frac{R_{\text{inst},i}(t, n_i)}{R_{\text{avg},i}(t)} \tag{2.8}$$

Weighted Proportional Fair (W-PF): In the W-PF algorithm, the term "weighted" indicates the consideration of user QoS weight. Overall, the metric is derived by considering the user QoS weight, user channel conditions and fairness. The QoS weight of a user is established by considering the QoS requirements of its radio bearers. The channel conditions of a user are determined by calculating the SINR of the user SRS carrying CSI received at the eNodeB. The receive SINR value is used to determine the instantaneous achievable bit rate for that user. The fairness criterion ensures that all users with pending uplink data get some share of resources

and user starvation is avoided. Mathematically, the TDPS metric value for user i is formulated as

$$\Lambda_i(t) = \frac{R_{\text{inst},i}(t, n_i)}{R_{\text{avg},i}(t)} \sum_a W_{i,a}(t) \tag{2.9}$$

where $W_{i,a}(t)$ is the QoS weight of the bearer a of user i at time t. $W_{i,a}(t)$ can be mathematically expressed as in (2.10).

$$W_{i,a}(t) = \frac{R_{\text{min},a}}{R_{\text{avg},i,a}(t)} \frac{\tau_{i,a}(t)}{\tau_{\text{max},a}} \rho_a(t) \tag{2.10}$$

where $R_{\text{min},a}$ is the bit rate budget (target data rate) and $\tau_{\text{max},a}$ is the end-to-end delay budget (target delay) of QoS class of bearer a, $R_{\text{avg},i,a}(t)$ is the average throughput of bearer a and $\tau_{i,a}(t)$ is the packet delay of bearer a of UE i, $\rho_a(t)$ is a variable with value set to 10 if $\tau_{i,a}(t)$ is above the threshold value of bearer a at time t (as the bearer packet delay has reached the threshold), otherwise equal to 1. The value 10 of $\rho_a(t)$ raises the metric value of bearer a by 10 times and ensures immediate scheduling. This value works well for the traffic load scenarios investigated in this thesis. The bit rate budget, packet delay budget and delay threshold values used in this thesis for the QoS classes VoIP, video, HTTP and FTP, are given in Table 2.1. Bit rate budget for a QoS class is defined according to its traffic models in this thesis. However, these values can be tuned by the network operators to modify the behavior of the scheduler.

Traffic type	**Bit rate budget (Kbps), $R_{\text{min},a}$**	**Packet end-to-end delay budget (ms), $\tau_{\text{max},a}$**	**Packet delay threshold (ms)**
VoIP	55	100	20
Video	132	300	100
HTTP	120	300	–
FTP	10	300	–

Table 2.1: Bearer bit rate budget, delay budget and delay threshold

2.1.4 Frequency Domain Packet Scheduler

The FDPS is the radio resource allocation phase of the uplink scheduling. The resources are allocated to users having high TDPS metric values. Low value users are discarded from the scheduling list in the current TTI. The uplink LTE system

bandwidth is divided into groups of PRBs that are allocated to a user as a whole unit. Such a group of PRBs is called a Resource Chunk (RC). Each RC consists of a fixed number of PRBs. The RC allocation is based on the FDPS metric values for each PRB of each RC of every user.

2.1.4.1 Frequency Domain Metric Algorithms

In frequency domain, the uplink scheduling algorithms are designed to map the users to the PRBs. Users are allocated a portion of bandwidth by the FDPS according to priority metric values determined by the FDPS metric algorithm. Therefore, the PRBs of all the users, allowed by TDPS to enter the FDPS, are given FDPS metric values. Some of the well-known FDPS metric algorithms and the Weighted Proportional Fair Scheduled (W-PFS) algorithm proposed in [Mar11] are explained here.

Maximum Throughput (MT): The frequency domain MT algorithm simply allocates a PRB to a user having maximum SINR value at a particular PRB. If $\lambda_{i,c}$ is the FDPS metric value of user i at PRB c, the metric value can be given as

$$\lambda_{i,c}(t) = SINR_{i,c}(t) \tag{2.11}$$

where $SINR_{i,c}(t)$ is the SINR value of user i at PRB c.

Proportionally Fair (PF): The frequency domain PF algorithm allocates the PRBs to users considering both the achievable throughput and fairness factors [MPKM08].

$$\lambda_{i,c}(t) = \frac{R_{\text{inst},i,c}(t)}{R_{\text{avg},i}(t)} \tag{2.12}$$

where $R_{\text{inst},i,c}$ is the instantaneously achievable throughput of user i at PRB c.

Proportional Fair Scheduled (PFS): The frequency domain PFS algorithm is a modified form of PF algorithm. The average throughput of a user over T TTIs is replaced by $R_{\text{sch,avg},i}(t)$ which is instantaneously achievable throughput of the user over only those previous T TTIs where i is scheduled. The PFS is usually preferred over PF as FDPS algorithm due to a biased behavior of PF towards higher GBR users as these users would get low priority due to this bias [ARC+08].

$$\lambda_{i,c}(t) = \frac{R_{\text{inst},i,c}(t)}{R_{\text{sch,avg},i}(t)} \tag{2.13}$$

GBR-aware Proportional Fair Scheduled (G-PFS): Frequency domain G-PFS is a modified form of PFS with the GBR weight playing a role in metric determination. This algorithm is used if the TDPS is expected not to provide fairness to the users having throughput requirements [ARC+08]. TDPS can only provide fairness if there is a sufficiently large number of users in the cell. However, in case of a small number of users, TDPS may not have any effect. In that case, the GBR-aware metric value can provide the required fairness. The G-PFS metric value can be expressed as

$$\lambda_{i,c}(t) = \frac{R_{\text{inst},i,c}(t)}{R_{\text{sch,avg},i}(t)} w_i(t) \tag{2.14}$$

where $w_i(t)$ is the FDPS weighting factor for users that are below their minimum required bit rates.

$$w_i(t) = \max(1, \frac{R_{\text{min},i}}{R_{\text{avg},i}(t)}) \tag{2.15}$$

Weighted Proportional Fair Scheduled (W-PFS): In the W-PFS algorithm, the QoS weight of the bearers of a user as in (2.10) is utilized. The metric is expressed mathematically as [McYZGTG12]

$$\lambda_{i,c}(t) = \frac{R_{\text{inst},i,c}(t)}{R_{\text{sch,avg},i}(t)} \sum_a W_{i,a}(t) \tag{2.16}$$

This algorithm considers the QoS limitations of the user for resource allocation.

2.1.4.2 RC Allocation Algorithm

In this thesis, a Fixed Size Chunk and Flexible Bandwidth (FSCFB) algorithm is proposed, which allocates RCs to users by finding out the RC combination which can achieve the best accumulative FDPS metric value; i.e. the highest possible sum of the RC metric value. In the process, the combinations that breach the contiguity constraint are discarded. Similarly, the combinations that allocate more PRBs to a user than its buffer requirement and the combinations which allocate more PRBs to a user than the limit indicated by PC are also discarded. The FSCFB is an adaptive transmission bandwidth algorithm, i.e., the users can obtain variable bandwidths. The algorithm procedures are summarized as follows:

1. Make a UE-RC table with each element as an RC metric value of a UE.

	RC_0	RC_1	RC_2
UE_0	7	6	13
UE_1	9	2	12

Table 2.2: A sample UE-RC table for two UEs and three RCs

2. Make all possible combinations of UE-RC allocation while satisfying the contiguity, buffer size and maximum allowed bandwidth constraints; determine the resulting global metric value for each combination.

3. Choose the combination with the best global metric value.

4. Obtain the desired resource allocation from the best combination.

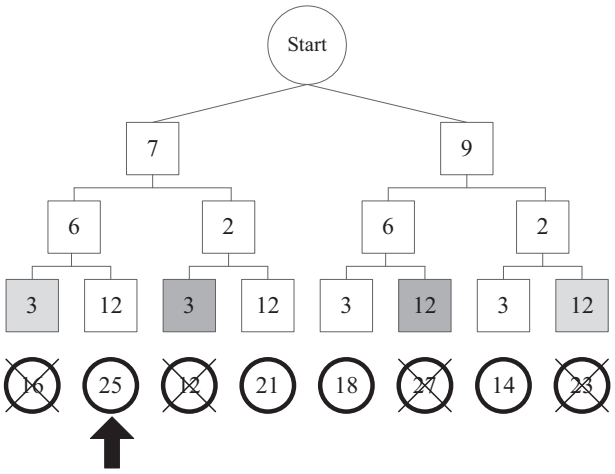

Figure 2.2: uDFS for two UEs and three RCs corresponding to Table 2.2

In step 1, a UE-RC table is generated as in Table 2.2. Each element in the table is an RC metric value of a UE, i.e. the sum of W-PFS metric values of the PRBs within the RC. Step 2 determines the possible allocation combinations. This is done using the unique Depth-First Search (uDFS) algorithm which is a search-tree based resource allocation algorithm proposed in this thesis, with contiguity, buffer size and maximum allowed bandwidth constraints. The uDFS is a sub-algorithm of FSCFB and it finds all possible combinations of RC allocation and discards the combinations that do not follow the constraints. Further depth of branches with a node breaching the constraints is not explored. The example of a uDFS tree given in Figure 2.2 corresponds to the UE-RC table in Table 2.2. This tree is based on the assumption that due to PC restrictions, only two RCs are allocated to a UE at most.

The dark grey colored nodes of the tree breach the contiguity constraint, whereas the light grey colored nodes breach the maximum PRBs constraint. In step 3, the surviving leaves (nodes without children) of the tree are checked to determine the global metric value for each combination. The combination giving the best global metric value represents the final RC allocation among the UEs in step 4.

2.1.5 Multi-bearer User Scheduling

Once the RCs are allocated, the users with multiple bearers divide the PRBs among the bearers. Each user bearer has its own QoS requirements. The user bearers require adequate resources to ensure QoS provision and avoid bearer starvation. In this thesis, the bearers receive weighted service, i.e. a bearer is served according to its QoS weight $W_{i,a}(t)$. However, the bearers having reached their packet delay threshold are given strict priority and the available resources are allocated to them before serving other bearers.

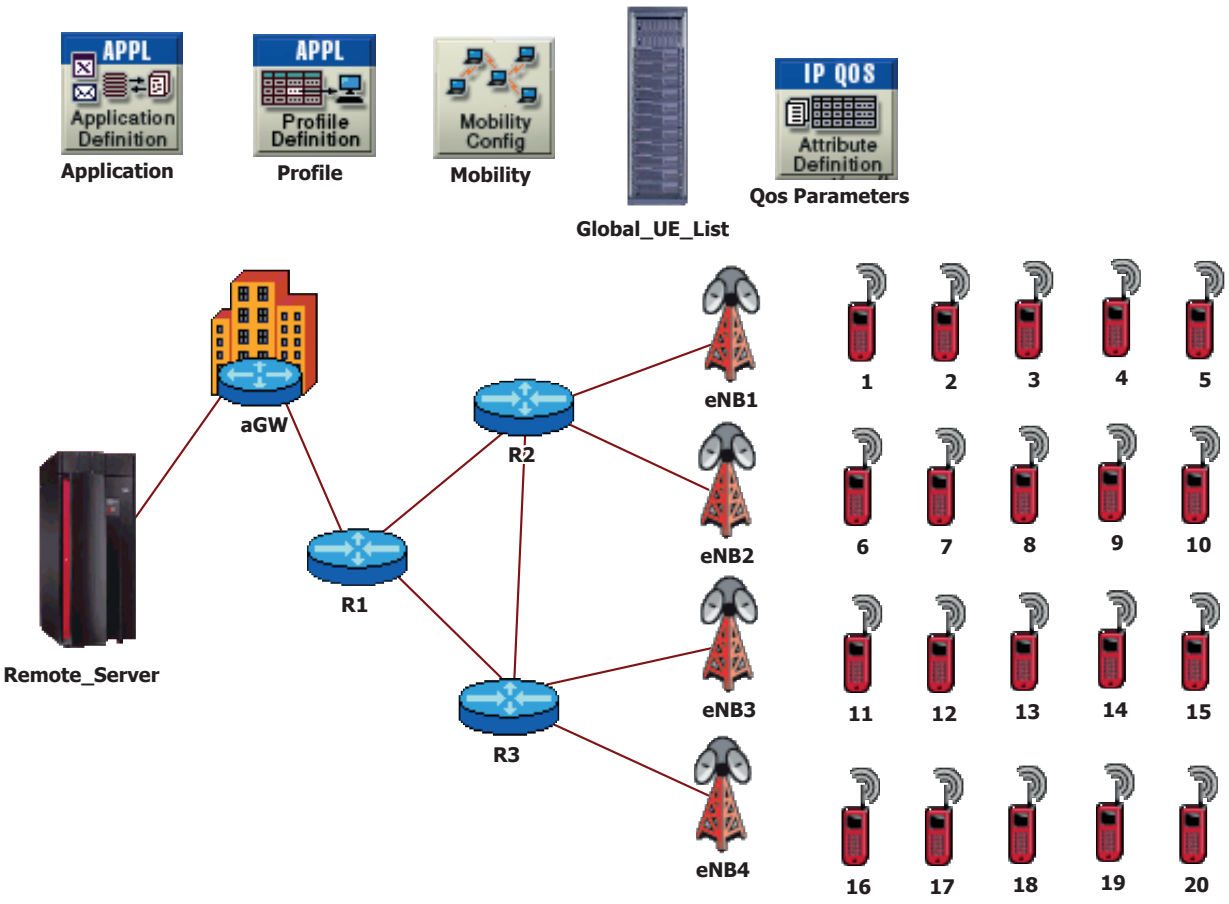

Figure 2.3: OPNET LTE simulation model

2.1.6 OPNET Modeler and Simulation Environment

The OPNET Modeler is used as the primary modeling, simulation and analysis tool of this thesis. The OPNET Modeler provides solutions for network research and development, as well as, application and network performance management [Riv14].

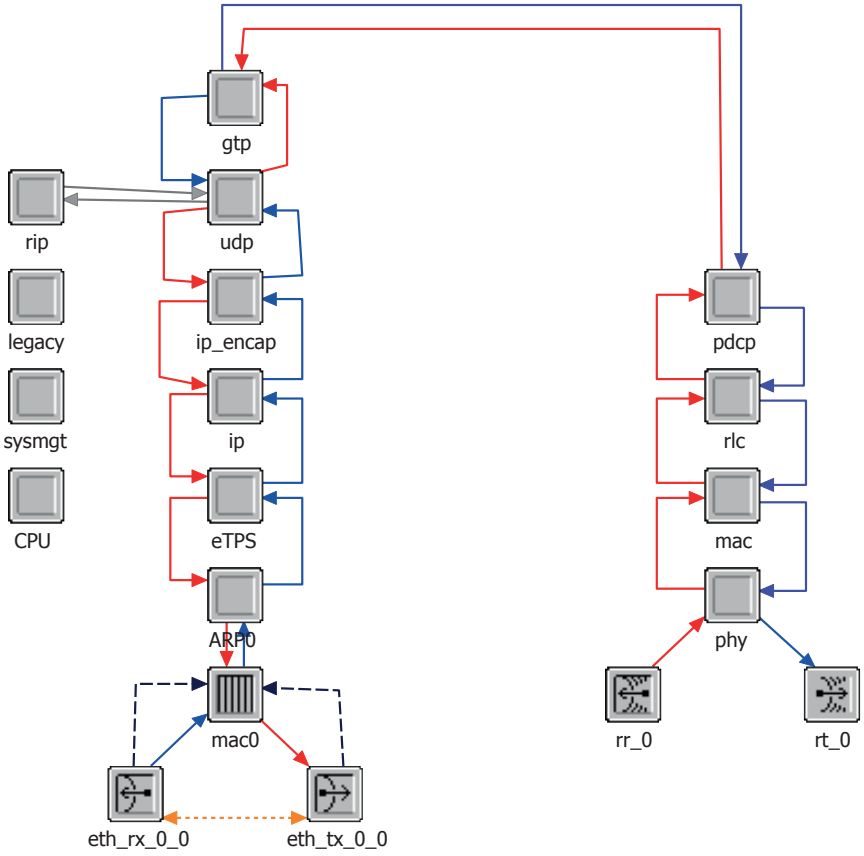

Figure 2.4: eNodeB node model

The LTE simulation model [ZWGTG11b] used in this thesis is depicted in Figure 2.3. This model consists of several nodes with LTE functionalities and protocols. The remote server node supports user applications in all the cells and also acts as a sink for the uplink data. The remote server is connected to the access GateWay (aGW) node via an Ethernet link of 20ms average delay. The aGW node consists of peer-to-peer protocols such as User Datagram Protocol (UDP), IP, and Ethernet towards the transport network; and also the protocols towards the remote server. The aGW and eNodeBs (eNB1, eNB2) are connected through the transport network of IP routers (R1, R2 and R3) configured according to the standard OPNET Differentiated Service (diffserv) model and routing protocols. The QoS parameters for the transport network are configured at the QoS parameters node

model for the traffic differentiation. The mobility node model consists of the mobility and channel models. The mobility model emulates the UE movement in a cell and updates the user location at every sampling interval of 1ms. The UE information is stored in the global user database (Global_UE_List) and is accessible from every node model in the system at any given time. The channel models include path loss, slow fading and fast fading. The development of the UE nodes follows the LTE Uu protocols and end-user application protocols. The Uu protocols are the PDCP, RLC, MAC and Physical (PHY) layers. The eNodeBs are developed according to the LTE Uu protocols and the transport related protocols (Figure 2.4). The PC functionality is in the PHY layer of UE nodes.

2.1.7 Simulation Parameters, Traffic Models and Results

The simulation parameter settings are given in Table 2.3. The simulation results (from [MWZ+14]) illustrate the performance of the BQA scheduler in comparison to BET, MT and PF schedulers in terms of throughput, QoS and fairness. The traffic models for VoIP, video, HTTP and FTP are also given in Table 2.3. The simulations are performed with loaded traffic environment in a single eNodeB having an omni-directional antenna. Therefore, no inter-eNodeB interference due to frequency reuse factor of 1 is considered in SINR calculations.

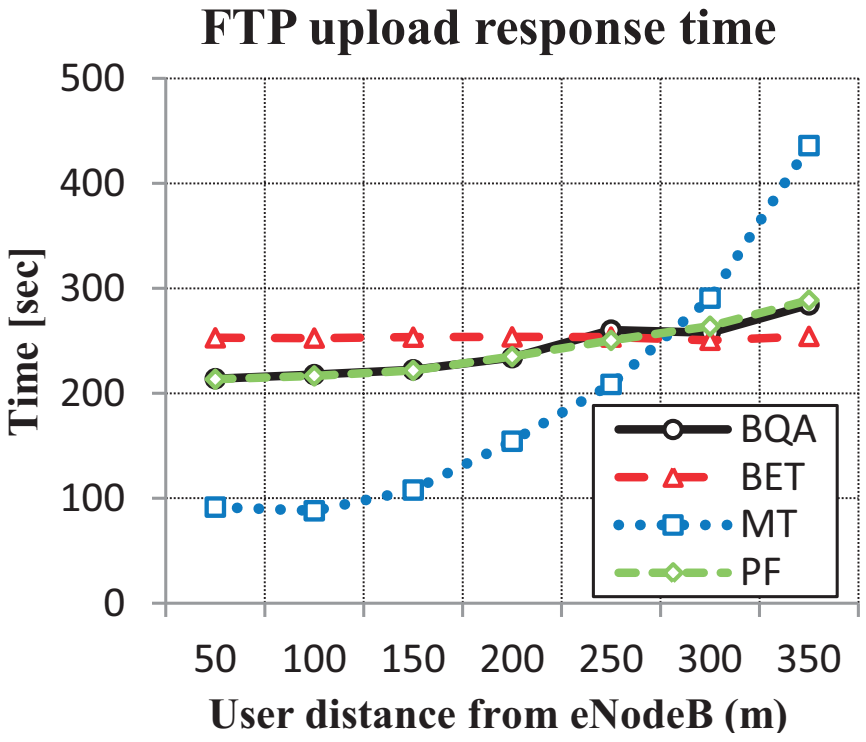

Figure 2.5: Average FTP upload response time under parameters in Table 2.3

Parameter	Setting
Cell layout	1 eNodeB, single cell
System bandwidth	25 PRBs ($\approx$ 5 MHz)
Cell radius	375 m
UE velocity	120 kmph
Max UE power	23 dBm
Path loss	128.1 + 37.6log10(R), R in km
Slow fading	Log-normal shadowing, 8 dB standard deviation, 0 dB mean, correlation 1
Fast fading	Jakes-like method [Cav00]
Noise per PRB	-120.447 dBm
Noise floor	9 dB
UE buffer size	Infinite
Power Control	FPC, $\alpha = 0.6, P_0 = -58$ dBm
Traffic environment	Loaded
Max FDPS UEs	5
RC size	5
VoIP traffic model	
Silence/talk spurt length	Exponential(3) sec
Encoder scheme	GSM EFR
Video traffic model	
Frame size	1200 bytes
Frame inter-arrival time	75 msec
HTTP traffic model	
Page size	100 Kbytes
Page inter-arrival time	12 sec
FTP traffic model	
File size	20 Mbytes
File inter-request time	Uniform distribution (min 80 sec, max 100) sec

Table 2.3: Main simulation parameters and traffic models

2.1.7.1 Fairness in Diverse Channel Conditions

The performance of the schedulers is compared in terms of fairness towards users with diverse channel conditions in this category of simulations. Seven users with uplink FTP traffic are placed at fixed locations in the cell such that the nearest user is placed at 50m from the eNodeB. The next user is located at a distance of 100m from the eNodeB and similarly, the rest of the users are positioned at distances with an increment of 50m each. The mobility is turned off. The UEs located in the vicinity of the eNodeB have good channel conditions and the UEs near the cell-

edge have poor channel conditions. The goal of these simulations is to analyze the scheduler performance towards good channel UEs and poor channel UEs. Ideally, cell edge UEs should also obtain adequate resources so that their QoS requirements are fulfilled.

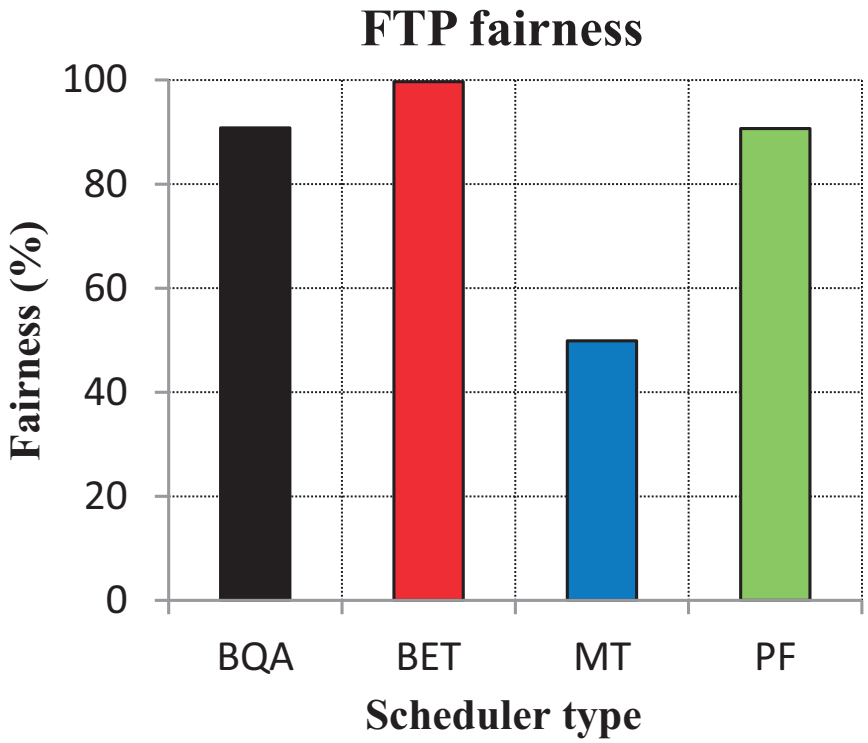

Figure 2.6: Fairness percentage comparison of schedulers (Table 2.3)

The results of these simulations are depicted in Figure 2.5 and Figure 2.6. In Figure 2.5, the user average FTP upload response time of an FTP file is depicted for all the UEs with various schedulers. The MT scheduler provides most radio resources to users with good channel conditions. Hence, the file upload time of the MT scheduler for the user at a distance of 350m is 300% higher than the user at 50m. As a result, the fairness percentage (determined using 2.17) of MT scheduler is below 50%. The BET scheduler provides resources to users equally without considering their channel conditions and, therefore, achieves a fairness percentage of 99.7%. The PF and the BQA schedulers maintain fairness by providing adequate resources to users with poor channel conditions by achieving 90.8 and 90.7% fairness percentage respectively. The fairness percentage comparison of the schedulers is depicted in Figure 2.6 and is expressed [ZZW+11] as:

$$Fairness\% = \left(1 - \frac{1}{M}\sum_{i=1}^{M}\frac{|R_i - R_{\text{avg}}|}{R_{\text{avg}}}\right) \times 100 \tag{2.17}$$

where R_i is the individual user Uu throughput, R_{avg} is the average of all user Uu throughputs and M is the number of users.

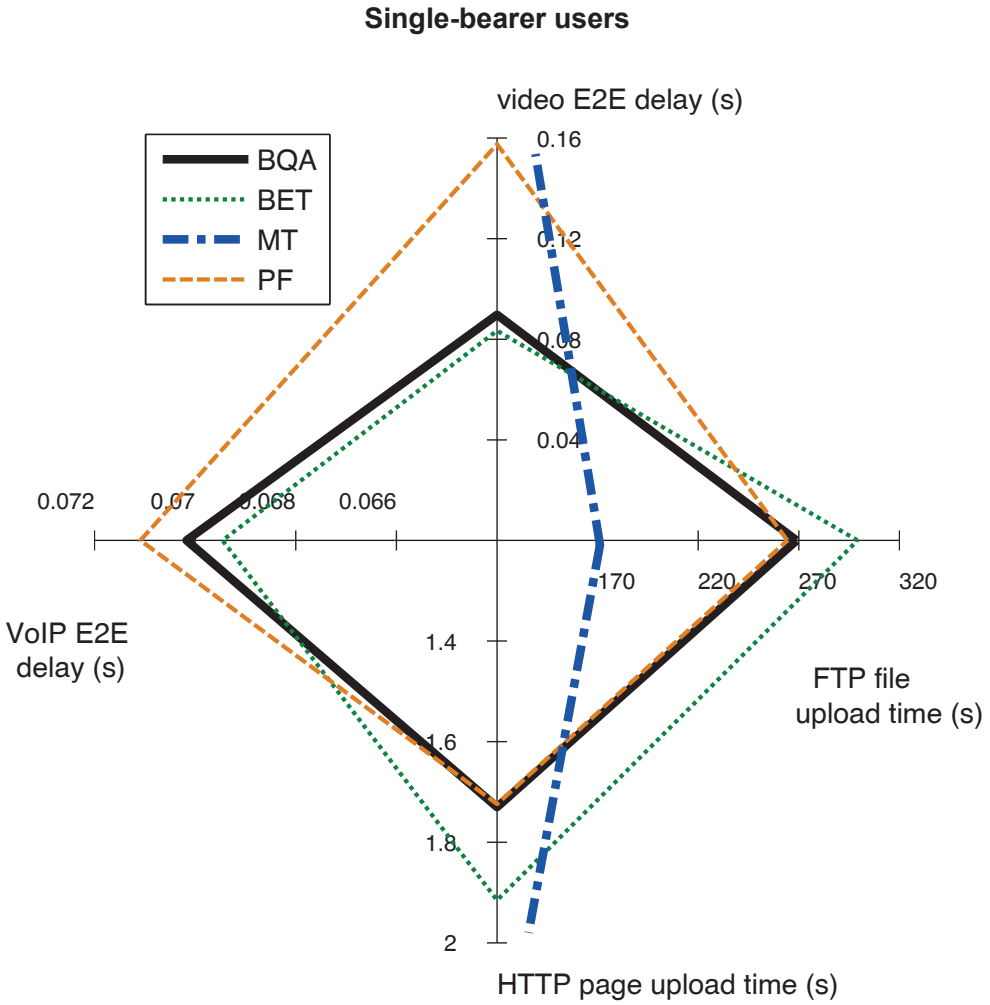

Figure 2.7: QoS performance comparison in single-bearer scenario (Table 2.3)

2.1.7.2 Performance in Single-Bearer Scenario

In the 2nd category of simulations, the performance of the schedulers is compared in terms of fairness and QoS provision to single-bearer UEs moving at a velocity of 120 km/h. In the cell; 8 VoIP, 8 video, 8 HTTP and 8 FTP users are deployed. The QoS performance of the schedulers is compared using spider web charts. In spider web charts, several QoS parameters can be compared in a single chart. The graphs in Figure 2.7 give an indication of the average time delay of the packets of the bearers. The FTP delay is represented by average file upload time, video and VoIP by average packet end-to-end delay, and HTTP by average page upload time.

The fairness and throughput performance comparison is illustrated in Figure 2.8. In Figure 2.7 and Figure 2.9, the delay performance of the schedulers is evaluated. Therefore, the smaller the scheduler graph size, the better the performance. In Figure 2.8 and Figure 2.10, the fairness and throughput performance is illustrated. Hence the larger the graph size, the better the performance.

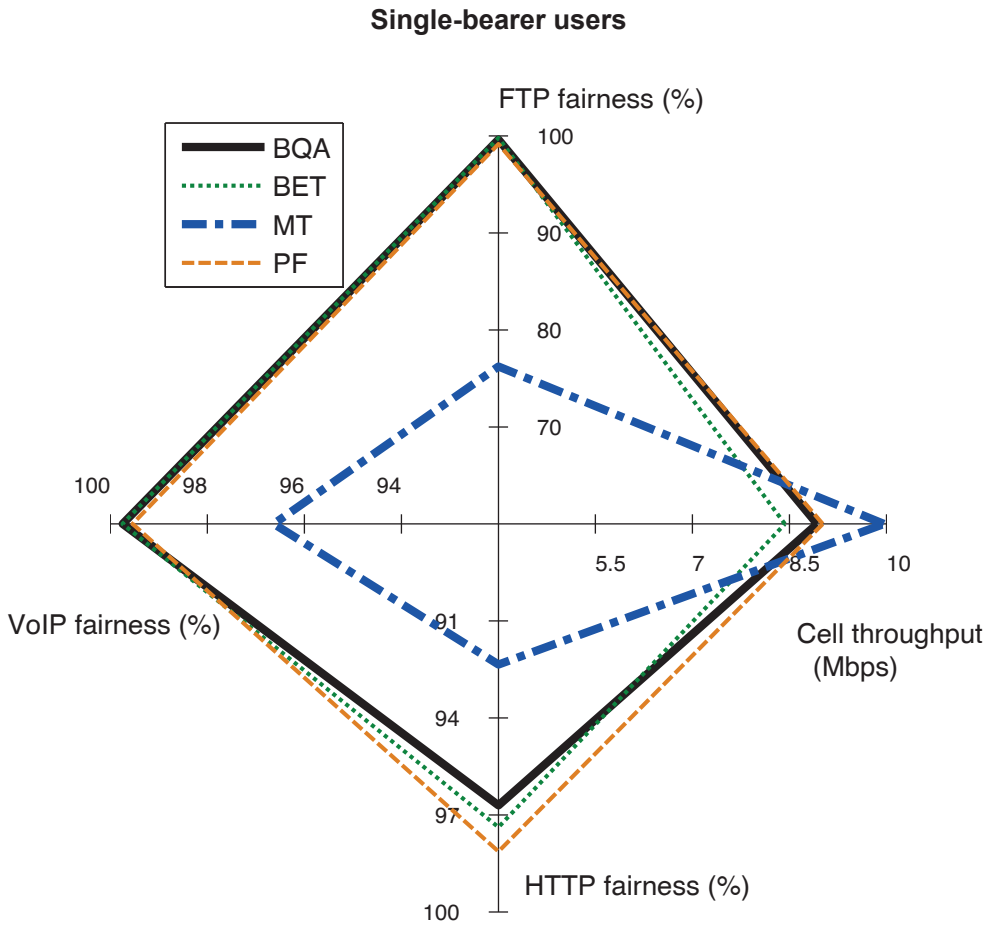

Figure 2.8: Throughput and fairness comparison in single-bearer scenario (Table 2.3)

Since the MT scheduler attempts to maximize the cell throughput, the FTP traffic is served better by the MT scheduler because the QoS unawareness of the scheduler means that delay sensitive traffic would not be prioritized. The only objective of the MT scheduler is to increase data rates. Hence, the MT scheduler serves FTP traffic nicely but other traffic types are not served appropriately. The delay results of other traffic types for the MT scheduler are huge and cannot be depicted in the scale of Figure 2.7. The BET scheduler strives for serving all the users fairly in terms of throughput. Resultantly, VoIP and video traffic with small message sizes are served with small average delay. However, the BET performance regarding FTP and HTTP bearers is worse. The PF scheduler maintains a balance between

maximizing cell throughput and fairness. PF results are between MT and BET results. The BQA scheduler attempts to both maximize throughput and provide fairness. Since, the BQA scheduler is QoS-aware as well, it gives higher priority to delay sensitive traffic. Therefore, the results in Figure 2.7 show that the performance of the BQA scheduler is far better than other schedulers with respect to delay sensitive traffic.

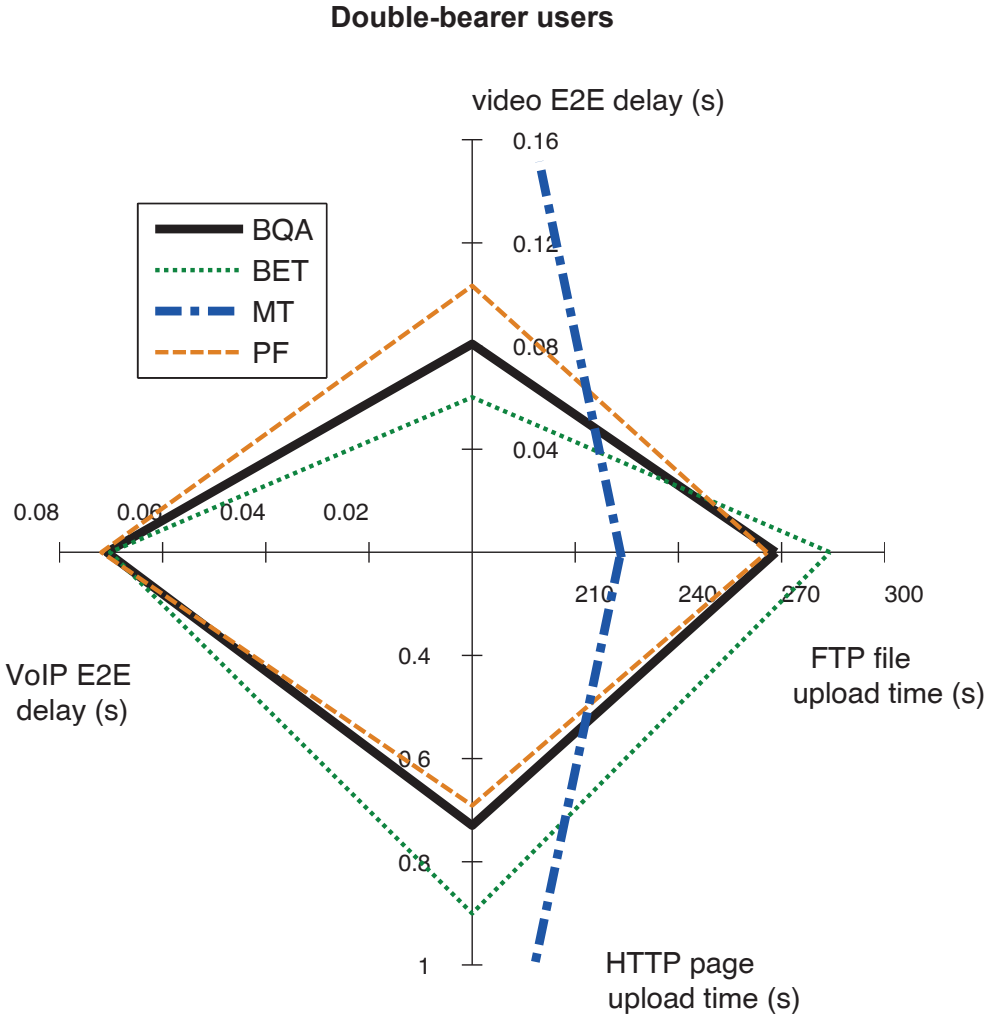

Figure 2.9: QoS performance comparison in double-bearer scenario (Table 2.3)

Figure 2.8 illustrates the comparison of schedulers in term of fairness for FTP, VoIP and HTTP traffic. The throughput performance of the schedulers is also compared in this graph. The performance of all the schedulers can be considered as fair, except the MT scheduler. The MT scheduler achieves maximum throughput without fairness. The BET scheduler achieves fairness but with low cell throughput. PF and BQA schedulers performance in these aspects is significantly similar, with BQA giving slightly better performance in terms of HTTP fairness.

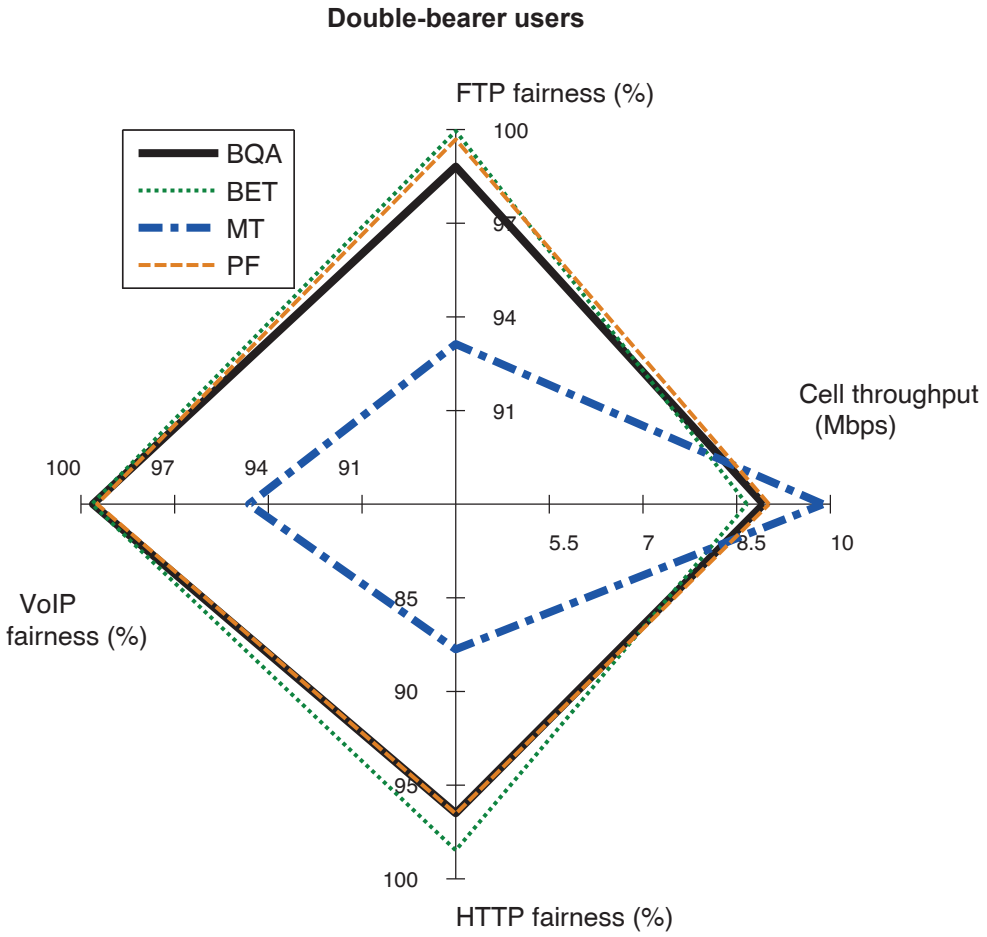

Figure 2.10: Throughput and fairness comparison in double-bearer scenario (Table 2.3)

2.1.7.3 Performance in Double-Bearer Scenario

In the 3rd simulations category, the performance of the schedulers is compared in terms of QoS provision to double-bearer UEs with a velocity of 120 km/h. In the cell of 16 users, 8 users have VoIP and HTTP bearers and 8 have video and FTP uplink data traffic. The QoS performance of the schedulers is compared in Figure 2.9, while the comparison of fairness and throughput performance is depicted in Figure 2.10. The results in the double-bearer scenario are similar to the single-bearer scenario results. The MT scheduler serves the FTP traffic better than the other schedulers but delay sensitive traffic is not served according to QoS requirements. The delay results of the MT scheduler for other traffic types are huge and not depicted in Figure 2.9. The BET scheduler fairly allocates resources and as a result, serves VoIP and video traffic with small average delay. However, the BET performance regarding FTP and HTTP bearers is not good. The PF scheduler results are, also in the double-bearer scenario, between MT and BET results.

The BQA scheduler gives higher priority to delay sensitive traffic. Therefore, the results in Figure 2.9 show that in the double-bearer scenario, the performance of the BQA scheduler is better than the other schedulers.

Figure 2.10 depicts the comparison of schedulers in terms of average cell throughput and fairness for FTP, VoIP and HTTP traffic. The fairness performance of all the schedulers is acceptable with the exception of the MT scheduler. The MT scheduler strives to maximize throughput without considering fairness. The BET scheduler achieves fairness but with inferior cell throughput performance. PF and BQA schedulers performance regarding fairness and throughput is quite similar in this scenario as well.

2.2 LTE-A Uplink Scheduling

The RRM in the LTE-A system is different from RRM in the LTE system. In LTE, only one component carrier can be used for data transmission for a UE. However, the LTE-A system is able to aggregate multiple carriers for data transmission. As explained in chapter 1, up to five CCs can be aggregated for a UE. In this thesis, a CCS scheme proposed in [WRP11] is reviewed, and a modification to the scheme is proposed. Based on the proposed CCS scheme, an LTE-A uplink scheduling scheme is presented.

2.2.1 Component Carrier Selection

The case of two intra-band contiguous CCs is considered in this thesis with one as a PCC and the other as an SCC. A simple CCS algorithm is that the users having a distance to the eNodeB larger than a threshold distance are assigned to a primary CC, otherwise both CCs. This algorithm is easy to implement and requires simple calculations; however, different threshold distances give different performances. If the best performance is to be achieved, the determination of the distance threshold requires a great amount of testing. Moreover, the network environment changes over time, while the distance threshold would not adapt accordingly, once it is set.

An effective path loss threshold ($L_{\text{threshold}}$) based CCS algorithm was proposed in [WRP11] to distinguish between power limited and non-power limited LTE-A stationary UEs.

$$L_{\text{threshold}} = L_{95\%} - \frac{10log_{10}N_{CC} + P_{\text{backoff}}}{\alpha} \tag{2.18}$$

where $L_{95\%}$ is the estimated 95 percentile user path loss, N_{CC} is the total number of

CCs and P_{backoff} is the estimated power backoff to model the effects of increased PAPR and cubic metric when a user transmits over multiple CCs simultaneously. With a higher power backoff, less LTE-A users can be assigned on multiple CCs due to the limitation of UE transmission power. The UEs are considered to be stationary in the cell in [WRP11]. If the path loss of the LTE-A UEs is higher than $L_{\text{threshold}}$, they are considered as power limited and assigned on a single CC; otherwise they are considered to be non-power limited and can use multiple CCs for data transmission. Hence, the cell-edge users would not experience any significant performance loss and would be scheduled over a single CC, while the non-power limited users can benefit from the advantages of a wider bandwidth.

This scheme is further extended in this thesis with improvements to achieve better system performance [MDL+14]. Instead of assuming that the UEs are stationary, a time variant radio channel model is used to get the real-time channel conditions of the UEs. Furthermore, not only the path loss is considered when determining the threshold and the number of CCs, the slow fading is also taken into consideration. Therefore, instead of labeling the UEs as power limited and non-power limited, the CCS would be based on the real-time channel conditions of the UEs. The proposed algorithm can be illustrated as:

$$L_{\text{threshold}} = (L+S)_{95\%} - \frac{10log_{10}N_{CC} + P_{\text{backoff}}}{\alpha} \tag{2.19}$$

At the time, when the sum of the UE's path loss and slow fading is higher than the threshold, one CC is assigned; otherwise, the user can use multiple CCs for data transmission.

2.2.2 Scheduler Overview

A major difference between LTE and LTE-A uplink scheduling is that the DFTS-OFDMA air interface of LTE-A allows the scheduling of resources without contiguity constraint. Moreover, the scheduler has to differentiate between the UEs on single and multiples CCs and allocate resources accordingly. The LTE-A uplink scheduler in this thesis is decoupled into TDPS and FDPS. The scheduler is denoted as Channel and QoS Aware (CQA) scheduling scheme.

2.2.3 Time Domain Packet Scheduler

The same W-PF algorithm of LTE is also used for TDPS in LTE-A uplink scheduling for each individual bearer to determine the TDPS metric value of the bearer. The bearers with high TDPS metric values are treated with high priority. However,

two separate lists of candidate bearers are created. One list consists of the GBR bearers, whereas the other list is for the non-GBR candidate bearers.

2.2.4 Frequency Domain Packet Scheduler

In the FDPS, the scheme in [ZWGTG11a] is modified in such a way that CCS is taken into account for scheduling decisions. Since GBR bearers require resources consistently, therefore the scheduling process starts with the candidate list of GBR bearers. The GBR candidate list is scheduled 1st by allocating a PRB to the top of the list bearer having highest priority before moving to the next bearer. This allocation process continues in a round robin fashion. After providing resources to GBR bearers, the non-GBR bearers of the non-GBR candidate list are allocated PRBs. The power limited users are restricted to PCC only. The CQA scheduling steps of FDPS for non-GBR users are summarized as follows:

1. At the beginning of each TTI, after the GBR list is scheduled, a certain number of high priority bearers are chosen from the non-GBR candidate list and placed in a subset candidate list. The bearer selection depends on bearer TDPS metric values.

2. The scheduler finds the PRB with highest SINR for the 1st bearer. This PRB is reserved for the 1st bearer and, therefore, the bearer's average SINR of the reserved PRBs is set as the effective SINR. The power limited users are able to obtain PRBs only from PCC. The MCS corresponding to the average SINR is considered for determining the achievable TBS.

3. The achievable TBS is compared to the buffer size of the bearer.
   a) If the achievable TBS is greater than the buffer size, the bearer can be completely served in this TTI, the PC constraint of the maximum PRBs for the UE is checked.
      i. If the number of PRBs reserved for the UE of this bearer exceeds the PC limit, the reserved PRB is not allocated and the UE is served with remaining reserved PRBs (if any) and discarded from the candidate list for this TTI. The reserved PRB is again available for other bearers.
      ii. Otherwise the reserved PRB is allocated to the bearer and scheduled.
   b) If the achievable TBS is smaller as compared to the buffer size, the bearer waits until the remaining bearers from the subset candidate list go through the above procedure.

4. Once all the bearers in the candidate list get a PRB reserved or get discarded after being scheduled, the above process is repeated again for the remaining bearers and the effective SINR of more than one PRB is calculated using AVI. This procedure is repeated until all the PRBs are allocated or all the bearers in the subset candidate list are served.

5. After the subset list is completely served, and still some bearers have to be served in the non-GBR candidate list, then in case of availability of PRBs, the next highest priority bearer in the candidate list is moved to the subset candidate list and provided PRBs according to the above procedure. This continues until there are no more PRBs available or no more bearers remaining to be served in candidate list.

Parameter	**Setting**
Number of CCs	2 (appr. 5 MHz each, intra-band contiguous)
System bandwidth	50 PRBs ($\approx$ 10 MHz)

Table 2.4: Simulation parameters in addition to Table 2.3

2.2.5 Simulation Parameters, Traffic Models and Results

The simulations are performed under the parameter settings and traffic models described in Table 2.3 with the addition of parameters given in Table 2.4. All the simulations analyze the performance of single-bearer UEs. The results are divided into two categories; the CCS scheme comparison results and the scheduler comparison results.

2.2.5.1 Component Carrier Selection Results

The proposed CCS scheme is analyzed and compared along with the algorithm proposed in [WRP11]. For comparison purposes, the CCS algorithm based on the distance between the UEs and the eNodeB is also evaluated as a reference scheme with cases of 100m and 200m considered. In case of 100m distance as the threshold, all the UEs having current location outside the 100m radius of the eNodeB are considered as power limited, whereas the UEs within the 100m radius are considered as non-power limited users. In the investigated scenarios, 40 voice, 40 video and 10 file transfer UEs are deployed in the cell. Figure 2.11 shows the throughput comparison among the three algorithms: path loss and slow fading based (PL_SF)

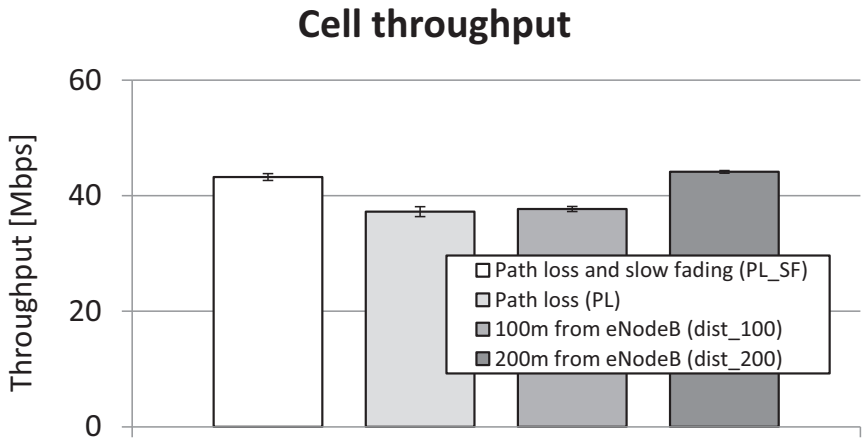

Figure 2.11: Comparison of cell throughput for CCS schemes

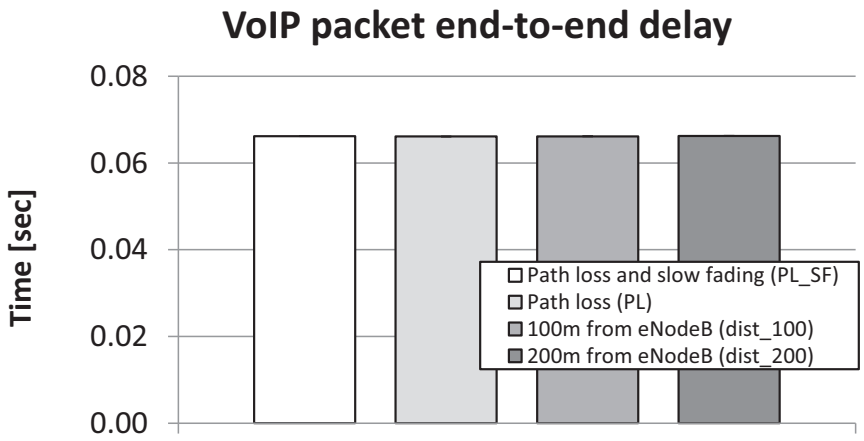

Figure 2.12: Comparison of VoIP packet end-to-end delay for CCS schemes

proposed in this thesis, path loss based (PL) [WRP11] and the one based on the distance to the eNodeB, using 100m and 200m as the thresholds. It is shown that the cell throughput is higher for the proposed path loss and slow fading based scheme (PL_SF) and the distance to the eNodeB of 200m as the threshold. Compared to the PL threshold scheme, the PL_SF gives a better performance, because in practice, both path loss and slow fading influence the transmitted signal. Threfore, the latter provides more accurate channel conditions and the CCS decision is more accurate than the PL scheme. For the scheme with the distance to the eNodeB as the threshold, varying distance thresholds would give different performances. In the cases considered here, the 200m threshold gives better performance as compared to 100m. However, this method is not adaptive and requires large effort on deciding the optimum distance threshold. Increasing the threshold distance allows more UEs to transmit over multiple CCs. However, if the threshold is increases beyond a

certain point, the cell edge users would also transmit over multiple CCs and result in interference to the neighboring cells as well as the loss of battery power.

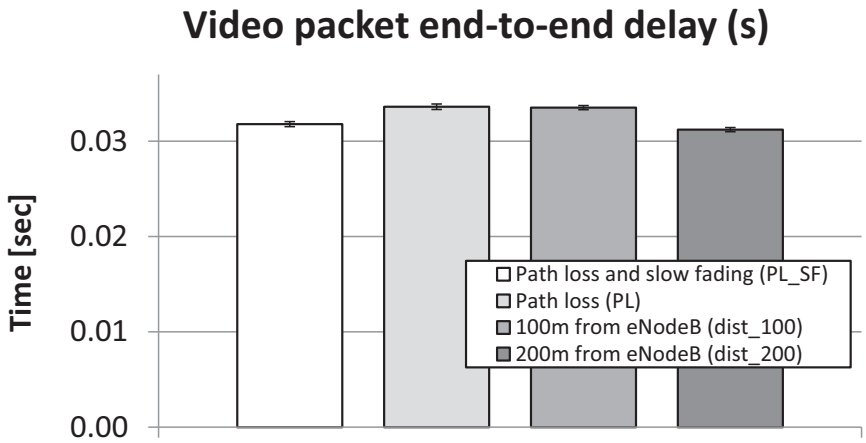

Figure 2.13: Comparison of video packet end-to-end delay for CCS schemes

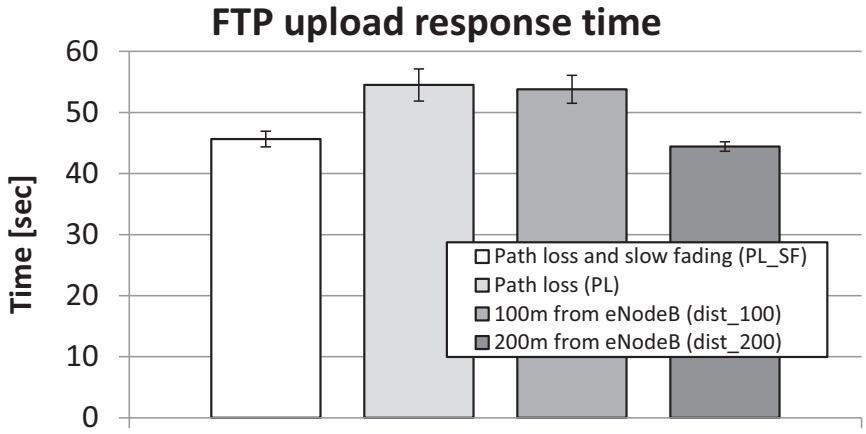

Figure 2.14: Comparison of file upload response time for CCS schemes

Figure 2.12 illustrates the average VoIP packet end-to-end delay. The VoIP users have GBR bearers and are served with highest priority. Therefore, the voice packet end-to-end delay does not vary much for different CCS algorithms. The video and the FTP users are non-GBR, so their performance varies with different CCS schemes. Figure 2.13 and Figure 2.14 present the average video end-to-end delay and average file upload time. The case with path loss and slow fading as the threshold, gives lower video delay and file upload time, compared to the one with only path loss as the threshold. Also, different distance thresholds result in good results for 200m and poor for 100m. The reason for the poor performance of users with lower threshold distance is the fact that all the users outside the 100m radius

are considered to be power limited. Therefore, a UE moving out of this radius is not able to transmit over multiple CCs.

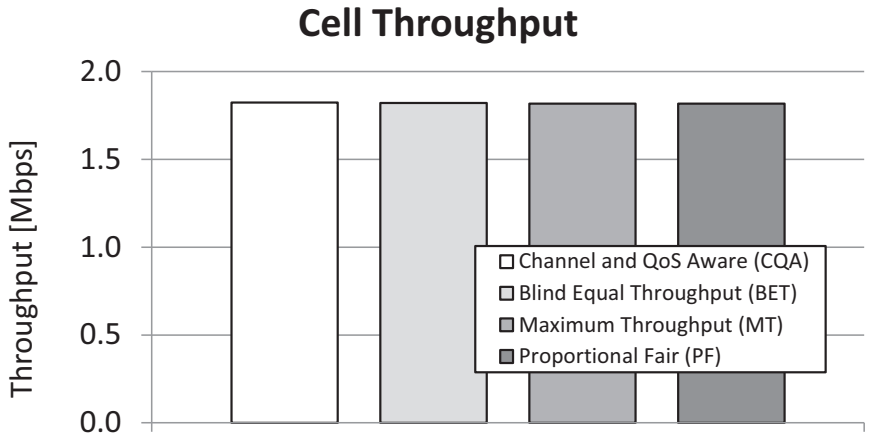

Figure 2.15: Comparison of cell throughput for various schedulers

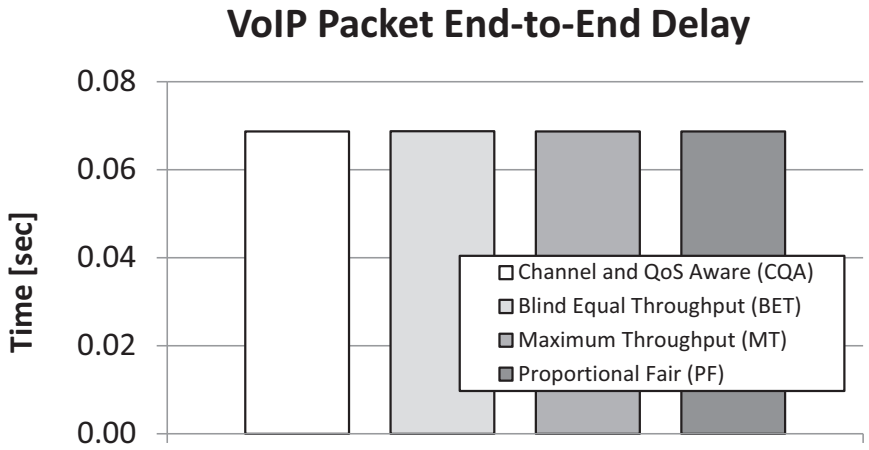

Figure 2.16: Comparison of VoIP end-to-end delay for various schedulers

2.2.5.2 Scheduling Results

For the comparison of QoS performance of various schedulers, the video traffic in the cell is reduced. Thus, instead of 40 video users, 10 video users are deployed in the cell. Figure 2.15 depicts the cell throughput comparison of the proposed LTE-A uplink scheduler and the reference schedulers BET, MT and PF schedulers. The comparison reveals that due to the less number of video user in the cell, the throughput performance of all the schedulers is quite comparable because all the schedulers are able to provide sufficient resources to users.

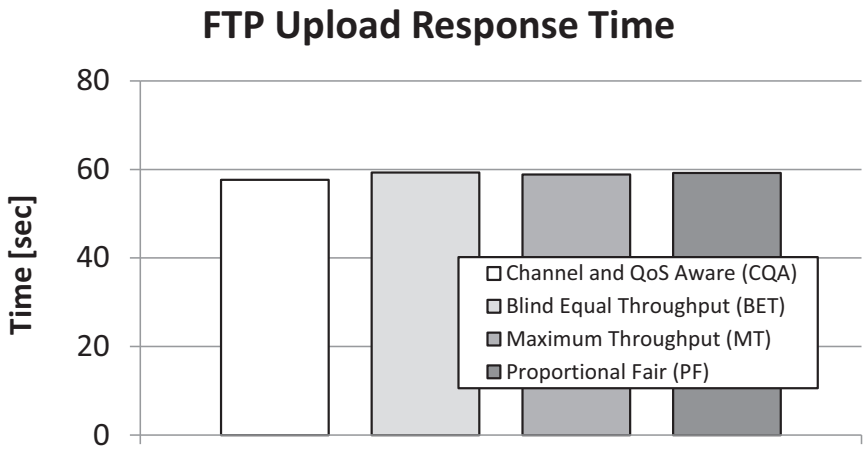

Figure 2.17: Comparison of file upload time for various schedulers

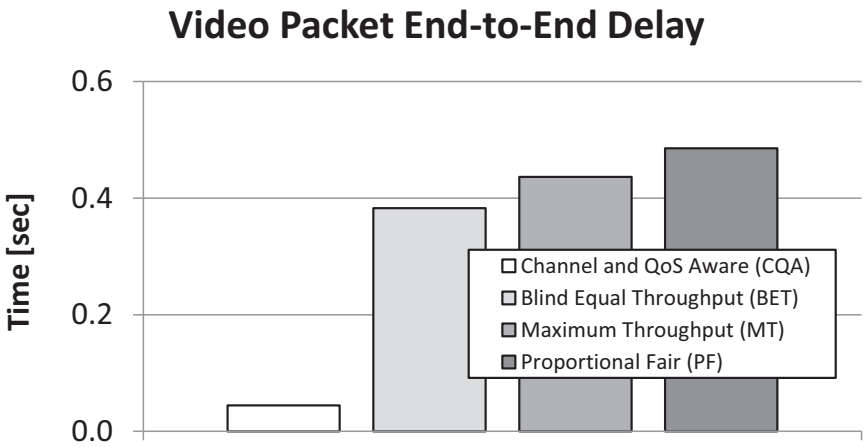

Figure 2.18: Comparison of video end-to-end delay for various schedulers

Figure 2.16 shows the average voice packet end-to-end delay. The performance of the VoIP users under various schedulers also does not vary significantly, since they are served with GBR at the highest priority. The file upload time (illustrated in Figure 2.17) with the proposed scheduler is the shortest, but the difference from other schedulers is not substantial. However, in Figure 2.18, the average video packet end-to-end delay results are quite contrasting. The proposed CQA scheduling scheme is designed with the awareness of the QoS requirements of the bearers; as a result the packet-end-to-end delay for video users is quite acceptable. But on the other hand, the BET scheduler, which provide equal throughput to all the users, neglects the delay requirements of video users. In case of the MT scheduler, the goal is to maximize throughput and the QoS requirements are not considered. The PF scheduler also ignores the QoS requirements of UE bearers. Hence, the end-to-end delay results for all the schedulers are very high. In summary, the pro-

posed CQA scheduler is able to treat traffic types according to their respective requirements due the QoS-awareness feature.

3 Machine-to-Machine Communication

Machine-to-Machine (M2M) communication is a global, rapidly-growing field of application and is of significant interest for many different devices and terminals. On the one hand, the variety and number of devices is growing quite swiftly; while on the other hand, the mobile data traffic is also expected to increase at a considerable pace in the near future [Cis14]. The term M2M communication refers to the terminals or devices, the so-called machines, with similar abilities to communicate with each other. The main idea behind M2M communication is that the importance of a machine grows when it is networked; whereas a network attains more importance when it interconnects more machines [Law04].

The "Internet of Things" is a new paradigm for the perception of getting more and more devices interconnected. The M2M communication can be regarded as a vital area in this field of communication. According to [CE11], the "Internet of Things" would expectedly extend the existing Internet by enabling the interconnection of a variety of devices. Similarly, the application domains would also benefit from this new concept, for example supervision of perishable goods in logistical processes.

The growing interest in remote monitoring applications such as health and environmental monitoring, as well as, the reduction in service and maintenance costs of cellular networks has changed the concept of the Internet [PMZG13]. Cellular broadband connectivity has become less expensive and global during recent years. The devices with integrated sensors, network interfaces, etc. are becoming more powerful, cheaper and smaller. This enables manufacturers to explore a variety of applications, devices, and new areas where those applications can be deployed.

3.1 M2M Network Architecture and Domains

The architecture of the M2M network can be divided into five elements according to ETSI [CWL12] as shown in Figure 3.1. These elements are:

1. Devices
2. Area networks

3. Gateway
4. Communication networks
5. Applications

3.1.1 Devices

The role of the devices in the M2M network architecture is to collect and transmit data such as sensors for pulse rate detection of a patient, temperature of a location, light intensity etc. Generally, these devices are connected to small local networks, for instance Wireless Local Area Network (WLAN), for the transmission of data to the M2M backend servers.

3.1.2 Area Networks

The M2M area networks establish the communication link between the devices and the gateway. These networks transmit data from the devices to the M2M gateway. There are several area networks used to link the devices to the gateway, for example WLAN, Bluetooth, ZigBee etc.

3.1.3 Gateway

The purpose of the gateway in the M2M architecture is to provide access to devices to another network for interconnection. Sensors deployed for the collection of information communicate with the network using M2M gateways.

3.1.4 Communication Networks

The primary goal of wired or wireless communication networks in the M2M architecture is to provide connection between the gateway and the applications. The well-known communication networks for M2M communication are LTE, LTE-A, WiMAX, satellite etc.

3.1.5 Applications

Data packets from communication networks are passed through several application services to be used by business agencies. Packets are usually analyzed by software agents.

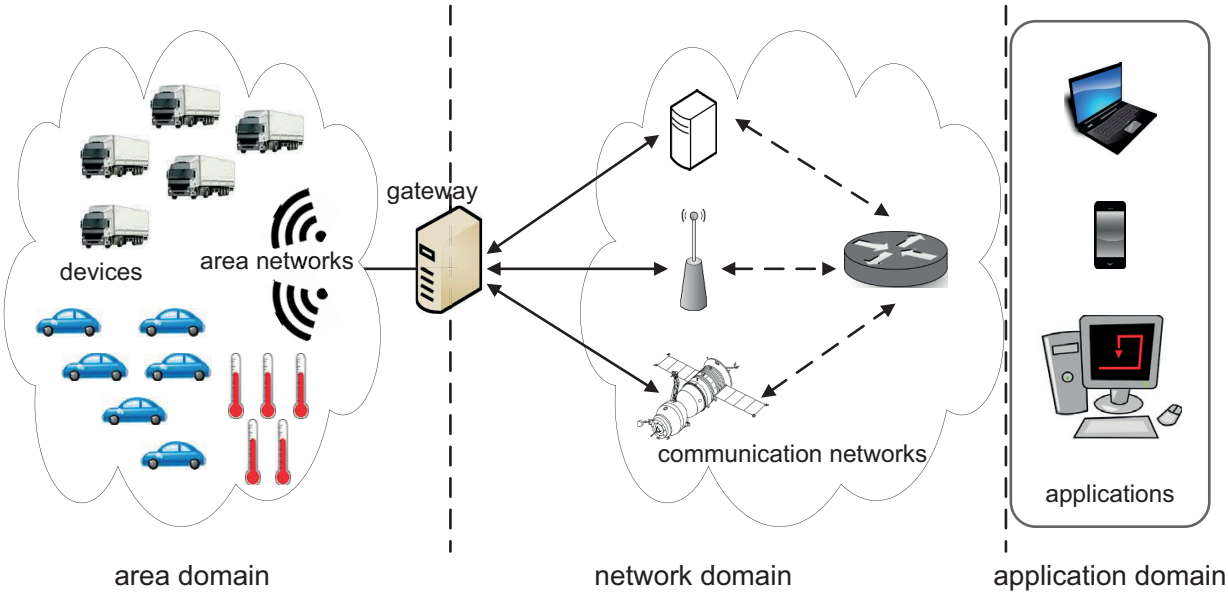

Figure 3.1: M2M network architecture [CWL12]

ETSI considers the technological standardization of M2M communication as a key factor for the long term development of the M2M technology. The five elements of the M2M architecture of ETSI are made up of three interlinked domains. The domains (as shown in Figure 3.1) are:

1. *Area domain* consists of the M2M devices, M2M area networks and the M2M gateway.

2. *Network domain* consists of all the wired or wireless networks of the element, communication networks of M2M architecture.

3. *Application domain* consists of all the application services.

3.2 M2M Standardization

M2M communication relies on several technologies spanning across various industries [CWL12]. As a result, the standardization of M2M communication is much more challenging than for the development of other conventional standards within a particular industry. Currently, the standards developing organizations such as 3GPP, ETSI, Institute of Electrical and Electronics Engineers (IEEE) and Telecommunications Industry Association (TIA) are working on M2M standardization. These standard developing organizations specify the network architectures and functions in accordance with the peculiar features of M2M communication.

An overview of the some of the application areas of standard developing organizations is provided next [WTJ+11].

3.2.1 3GPP

The 3GPP specifies features and requirements for M2M communication [23.12] with machines connected in such a way that the devices communicate either with other devices or with (one or more) servers. The 3GPP faces the issue of the scarcity of IP addresses and the congestion of control and data signals for a large number of M2M devices. Some of these issues have been addressed by proposals such as using IPv6 address, dual-stack address and group-based management for M2M devices.

3.2.2 ETSI

The primary goal of ETSI is to standardize service capabilities to provide common functions to contrasting M2M applications [10213]. ETSI also focuses on integrating M2M traffic to the advanced 3GPP wireless communication networks. ETSI covers M2M application areas like E-healthcare, smart grid, smart metering, connected consumers etc.

3.2.3 IEEE

IEEE standards mainly cover the air interface and related functions associated with the wireless connectivity of M2M devices [Cho11]. The optimization of air interface for mass device communication, usage of sub-GHz spectrum and smart grid are among the main M2M issues for IEEE.

3.2.4 oneM2M

The main purpose of one M2M is developing technical specifications to address the importance of an easily embeddable common M2M service layer for different softwares and hardwares [one14]. This can help in globally connecting devices having various traffic types and related to several application areas. At the moment, 227 partners and members are working on M2M specifications under the oneM2M umbrella. Five working groups of oneM2M are working on (1) requirements, (2) architecture, (3) protocols, (4) security and (5) management, abstraction and semantics.

3.3 M2M Application Areas

M2M communication can be deployed in several areas. The implementation areas include logistics, traffic systems, smart metering and monitoring, E-healthcare etc. M2M communications are expected to have a strong impact on these implementation areas and offer extensive improvements.

3.3.1 Logistics

M2M communications are expected to have a great impact towards improving the standards of logistic processes. The storage period of goods starts from departure at the source to arrival at the destination. In this period, the monitoring of several environmental factors including temperature, pressure, humidity, light intensity and location etc. is required. Goods are usually sensitive to environmental changes and regular monitoring is necessary. The sensor readings can be transmitted using M2M communications to take appropriate restorative measures and avoid unwanted situations. Therefore, M2M communications can bring extensive improvement to the logistic processes.

In the logistics industry, the delivery of goods with assurance of time, place, quantity and quality is of immense importance [SRMG09]. The hurdles in fulfilling these guarantees are caused by the increasing requirements and expectations of the consumers from the limited available resources. Such changes in the environment are the causes of "dynamics in logistics". Consequences of these environmental changes are re-routing, alteration of production strategies, shifting bottlenecks etc. In order to minimize the adverse effects of these dynamics, business groups continually strive for devising remedial measures. M2M communications can drastically enhance such measures if the mobile networks are designed in a manner that they are capable of handling such communications.

Fleet management is also a major application of M2M communication in logistics [Vod14]. The location, speed, fuel consumption etc. of vehicles and containers can be tracked regularly. The collected information is used to evaluate the transportation process for achieving in time delivery, avoiding routes not feasible for shipment and re-routing if required.

3.3.2 Smart Metering and Monitoring

The M2M application area of smart metering and monitoring is considered to be a major motivational force behind the rise of M2M communication [NXW11]. The primary goal of M2M smart sensors is to monitor the energy consumption at residential areas and commercial workplaces. The M2M devices perform the tasks of

acquisition of energy consumption data, processing of data and regular feedback of consumption data to the customers etc. The data obtained by the deployed devices for metering or monitoring purposes is sent to the backend servers. Detection of excessive energy consumption can trigger the transmission of warning messages to the consumer. This detection can be used to reduce electricity, gas, water, heating, cooling etc. costs for private public and enterprises.

The M2M application area of smart metering and monitoring encompasses several concepts [Voj08]. The concept of advanced metering infrastructure has been developed to control and monitor the energy consuming appliances inside customer premises. The smart pricing concept involves real-time pricing, time-variable pricing and differentiated pricing. These concepts are utilized by the smart devices and energy management systems which are able to make intelligent decisions based on smart prices. The implementation of these concept is quite challenging due to the fact that it requires enabling communications between millions of power meters [LTCH09].

3.3.3 Intelligent Traffic Systems

Intelligent Traffic Systems (ITS) is the application area of M2M communication where methods and technologies are developed for enabling communication of ITS applications with vehicles and traffic infrastructure [fD11]. Parameters from automotive industries such as transportation time, traffic collision avoidance, on-board safety, fuel consumption, parking time etc. are managed by the ITS applications. Communication of a vehicle with other vehicles as well as with ITS infrastructure is the primary objective of ITS.

In ITS, remote monitoring of vehicular data can be performed by the connecting M2M devices in the vehicle to the ITS application server. At the server side, the mileage of the vehicle can be monitored and an alarm can be triggered if mileage limit for the safety inspection is reached. Similarly, the temperature of the engine can be measured and alerts can be issued in case of fire along with the location of the vehicle. Vehicular health can also be diagnosed with the help of ITS by using M2M devices in the car for transmitting health data such as light bulbs state, cable connections state, fuse state etc. Monitoring the railway tracks to check obstacles on the track can help in accident avoidance. In the same way, the brake state, engine state, wheel state can also be monitored. With ITS, vehicle collision management can be enhanced by sending sensor measurements of the gap between vehicles to the infrastructure. In case of a collision, the other vehicles in the vicinity can be informed about the location of the event.

3.3.4 E-healthcare

The goal of e-healthcare is to establish a relationship between a patient and a healthcare organization using the M2M communication [fD11]. E-healthcare is defined as "the application of Internet and other related technologies in the healthcare industry to improve the access, efficiency, effectiveness, and quality of clinical and business processes utilized by healthcare organizations, practitioners, patients, and consumers to improve the health status of patients" [Mar02]. The emergence of wireless networks has developed the concepts of providing healthcare and patient monitoring facilities with mobility, even outside the premises of hospitals and dispensaries [MPZ+13]. Generally, the health condition of a patient requiring frequent check-ups is monitored on the hospital bed. The patient health parameters like temperature, blood pressure, heartbeat rate, etc. are monitored by the doctors and nursing staff. The hospitals, beds and staff are scarce and valuable public resources. The management of such resources is a challenging task for healthcare organizations even in advanced countries with abundance of resources.

As a result of the current technological advancements, the notion of monitoring patient health status from remote locations has been investigated by employing wireless networks [NH12] [KY08]. The condition of patients can be monitored while they are at home wearing sensors connected to a capillary network. The health parameters such as temperature, electrocardiography data, or electronic images of the patient can be sent to the hospital from a remote location by wrapping a sensor belt around the chest. In this way, convenience can be provided to the healthcare organizations and to the patients by decreasing the waiting time in the hospitals quite considerably. Additionally, the daily routine activities of the patients are not disturbed while staying at home. In case of sensors detecting unusual health status of the patient, alerts can be sent to the healthcare organization. The healthcare organizations are required to react to alerts only and not when the health status is usual. Moreover, if the patient stays away from the hospital environment, the risks of infections and other diseases can be avoided. However, a restriction on the patient is that capillary networks are helpful only inside buildings. Capillary networks are not a solution for outdoor mobility, for example the patient is walking down a street or driving a car. In such cases, the cellular networks can play a vital role in providing e-healthcare services.

3.4 M2M Traffic

In contrast to the traditional end-to-end communications involving human intervention, M2M devices transfer data without direct (or with minimum) human in-

tervention [WPSM04]. This results in traffic generation and usage behavior being different from the traditional human based traffic. The current M2M communications are based on the prevailing wireless communication standard technologies like GSM. These contemporary systems are sufficient for the current M2M traffic demands. These systems also offer low-cost and convenient deployment of M2M devices along with roaming facilities.

3.4.1 Traffic Trends

The characteristics of M2M traffic are quite different from the existing human based traditional network traffic. The traditional traffic follows a certain session length, data volume and interaction frequency. However, M2M traffic pattern are different from traditional traffic. One of the most significant differences is the amount of data per transmitted packet. In most of the applications, the packet size is usually very small, i.e. only a few bytes. In such applications, the M2M devices report sensor data, such as temperature, humidity, etc. The transmitted packets consist of the measurement data and the corresponding protocol overhead. Generally, the overhead is kept as small as possible, whereas the actual payload of measurement data differs according to applications.

In general, an M2M device reports the measurement data in periodical intervals to a remote end-point. Although the duration of the interval might range from minutes to hours [SJL+12], the multitude of distributed M2M devices in a cellular network might create a considerably dense scenario. In the meantime, the UEs having traditional traffic with network connectivity are also present. According to [Cis14], M2M traffic is expected to grow 11-fold in the forecast period of between 2013 and 2018, which corresponds to an annual growth of 61%. It is estimated that the number of devices with mobile connectivity might exceed the population of the world. Although smartphones and personal computers including laptops and notebooks would be the most dominant devices in terms of data traffic generation, but the M2M devices would contribute about 19.7% of the total mobile data traffic in 2018 as compared to 4.9% in 2013. The per-M2M-device usage growth in megabytes per month would rise from 61 to 451 in the forecast period.

3.4.2 M2M Issues

The probing issues that M2M communications face are related to the distinctive characteristics of M2M traffic. The most challenging issue is the expectedly large number of M2M messages. As explained above, a huge number of M2M devices are expected to be deployed in the future for M2M communications. For example,

the logistic processes would be supported by M2M devices deployed in transport vehicles, ships, factories, and storage houses. This forecast requires that the mobile communication standard committees revise the networking designs for efficient support of these devices. Another major concern would be the varying size of the M2M messages. The size of a M2M message could be just a few bits as in case of a simple temperature reading in a container. Similarly, messages of the order of megabytes could also be transmitted as in case of a video monitoring device.

3.5 Impact of M2M Traffic on LTE and LTE-A Performance

User mobility is a very special feature of cellular networks. Support for mobility is a major requirement of several M2M applications. In e-healthcare, logistics and ITS, it is required that the connectivity is provided even when the devices are moving at high velocities. In case of e-healthcare, the patients are often advised to perform outdoor physical activities like jogging, walking, exercises, etc. To carry out such activities, the patients are required to leave their home (or office). In such cases, the home WLAN cannot be a viable solution for e-healthcare services provision. The same is the situation when M2M devices are have high mobility in logistics and ITS scenarios. Hence, the significant role of mobile networks is expected to grow in future.

As explained, M2M communications are presently based on contemporary mobile communication technologies like GSM, which are fulfilling the requirements of certain M2M applications sufficiently. However, it is expected that the number of devices based on M2M communications might undergo an exponential growth. LTE and LTE-A are the expected future technologies for providing M2M services. It is anticipated that the M2M applications would offer a diverse range of services, including narrowband applications. However, the main problem with LTE and LTE-A is that these standards have been primarily developed for broadband data services. As a result, these standards may not achieve spectrum and cost efficiency for narrowband applications. The integration of M2M traffic, having low data rates and small packet sizes, might have an adverse impact on the performance of LTE/LTE-A systems.

As described, the smallest resource unit in LTE/LTE-A systems, that can be allocated to a device is the PRB. When channel conditions of a device are favorable, one PRB is capable of transmitting several kilobytes of data. So, if a single PRB is allocated to an M2M device, the impact on spectral efficiency might be negative. Transmission of synchronous messages to the network is also a concern and can drastically degrade the network performance and block resources for other devices.

This kind of situation can arise in cases of emergencies, such as fire, flood etc. For instance, in case of fire in a warehouse, the alarms inside the building and vehicles would simultaneously trigger resource requests. Due to diverse M2M applications expected to emerge in the near future, the issue of fulfilling the varying QoS requirements (regarding delay, throughput and packet loss rate) of various types of M2M messages is a challenging problem and needs to be addressed. Designing low complexity algorithms would also be a strenuous task. All these challenges to support M2M communications with LTE/LTE-A networks should be addressed in such a manner that the performance of traditional traffic users are not hindered.

3.6 Simulation Parameters, Traffic Models and Results

In order to illustrate the impact of M2M traffic on LTE system QoS performance, the OPNET simulation model explained in section 2.1.6 is used. The simulations are performed in such a way that, along with regular LTE uplink traffic like VoIP, video and FTP, the M2M traffic is also simulated.

3.6.1 Logistics

The M2M traffic is based on logistics data traffic given in [PMZG13] (Table 3.1). The parameter settings and the traffic models are given in Table 3.1. Five simulation scenarios are considered for evaluation where the regular LTE traffic load is kept constant and the M2M traffic load is increased in subsequent scenarios. In all scenarios, the number of VoIP, video, and FTP users is kept 10 based on the traffic models in Table 3.1. However, the number of M2M devices communicating with the eNodeB in the scenarios is varied in such a way that in the 1^{st} scenario, no M2M devices are deployed in the cell. In the 2^{nd} scenario, 300 devices are simulated in the cell. The number of devices is increased by 150 in each of the subsequent scenarios. Hence, the number of M2M devices in the scenarios is 0, 300, 450, 600 and 750 [PMZG13]. The uplink scheduler at the eNodeB gives higher priority to the VoIP and video data traffic as compared to the M2M traffic. However, the FTP traffic is given lower priority as compared to the M2M traffic.

The simulations results of LTE regular traffic users are compared using spider web chart in Figure 3.2. The QoS performance of VoIP and video users in terms of packet end-to-end delay shows that increasing the M2M traffic load has no significant impact on the performance of these users. The average packet end-to-end delay for VoIP and video users remain almost constant in all the scenarios. The reason for this behavior is that the LTE scheduler in [MWZ+14] is used here and it gives priority to delay sensitive traffic of VoIP and video users. On the other hand,

Parameter	**Setting**
Cell layout	1 eNodeB, single cell
System bandwidth	25 PRBs ($\approx$ 5 MHz)
Cell radius	375 m
UE velocity	120 kmph
Max UE power	23 dBm
Path loss	128.1 + 37.6log10(R), R in km
Slow fading	Log-normal shadowing, 8 dB standard deviation, 0 dB mean, correlation 1
Fast fading	Jakes-like method [Cav00]
UE buffer size	Infinite
Power Control	FPC, $\alpha = 0.6, P_0 = -58$ dBm
Traffic environment	Loaded
Uplink scheduler	BQA [MWZ+14]
VoIP traffic model	
Silence/talk spurt length	Exponential(3) sec
Encoder scheme	GSM EFR
Video traffic model	
Frame size	Constant(1,200) byte
Frame inter-arrival time	75 msec
FTP traffic model	
File size	Constant(20) Mbyte
File inter-request time	Uniform (min 80, max 100) sec
HTTP traffic model	
Page size	Constant(100) Kbyte
Page inter-request time	Constant (12) sec
M2M traffic model for logistics	
Message size	Constant(6) Kbyte
Message inter-transmission time	Constant(60) sec
M2M traffic model for e-healthcare	
Message size	Constant(1200) byte
Message inter-transmission time	Constant(1) sec

Table 3.1: Main simulation parameters and traffic models

if the performance of best effort traffic users with FTP data is considered, it can be clearly seen that increasing the number of M2M devices causes higher average file upload response times. Even in the scenario with only 300 M2M devices, the QoS performance degradation of LTE traffic is noticable. The average delay of FTP file upload response time increases from 231.368 sec to 249.611 sec, although M2M traffic in the cell coverage area is very little as compared to the expected mass

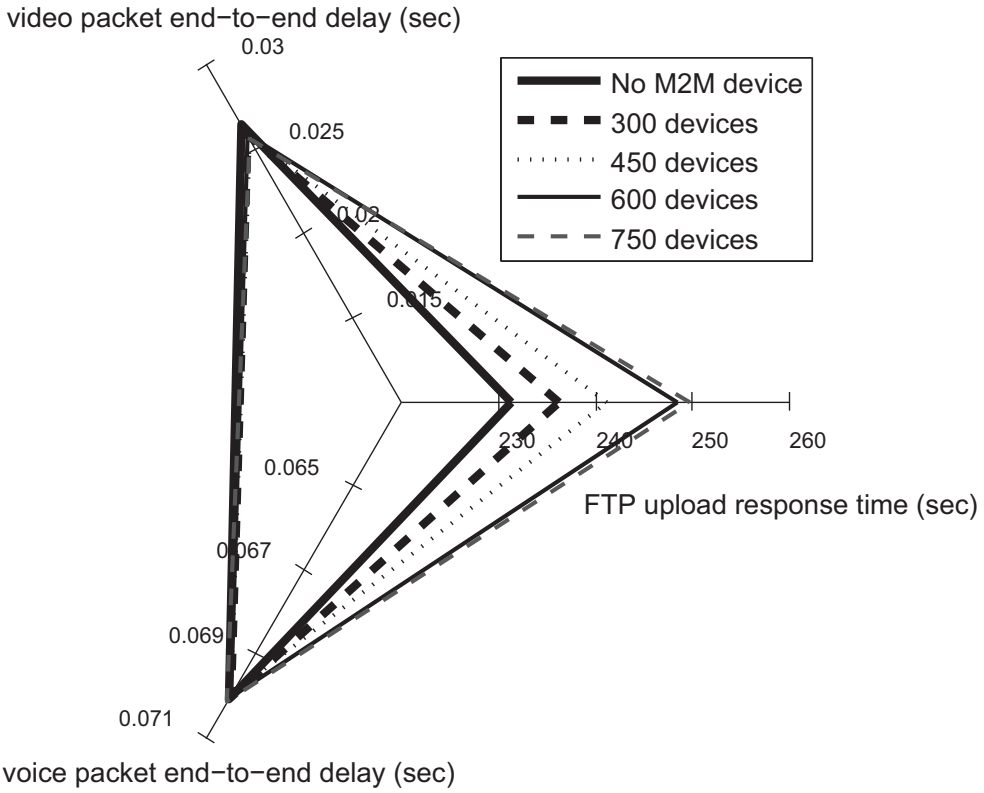

Figure 3.2: Cell QoS performance with logistics data traffic [PMZG13]

M2M deployment in future.

3.6.2 E-healthcare

The impact of M2M traffic on the QoS and throughput performance of regular LTE uplink traffic for e-healthcare scenarios is evaluated in this subsection. The regular LTE uplink traffic consists of four types of traffic, i.e., VoIP, video, HTTP (such as a mobile point-of-sale) and FTP. The M2M traffic model for e-healthcare is based on the model in [MPZ+13] and given in Table 3.1. The uplink HTTP traffic is similar to web browsing traffic, but web browsing data is generally in downlink direction. The voice and e-healthcare traffic is treated as the highest priority traffic by the scheduler. Video, HTTP and FTP traffic have lower priorities than the M2M traffic. In all scenarios, the regular LTE traffic load remains unchanged. However, the M2M data traffic load is varied from scenario to scenario. The number of VoIP, video, HTTP and FTP users is kept 8 each in all the scenarios. There is no M2M traffic in the 1st scenario. In the subsequent scenarios, the number of M2M devices is increased by 50.

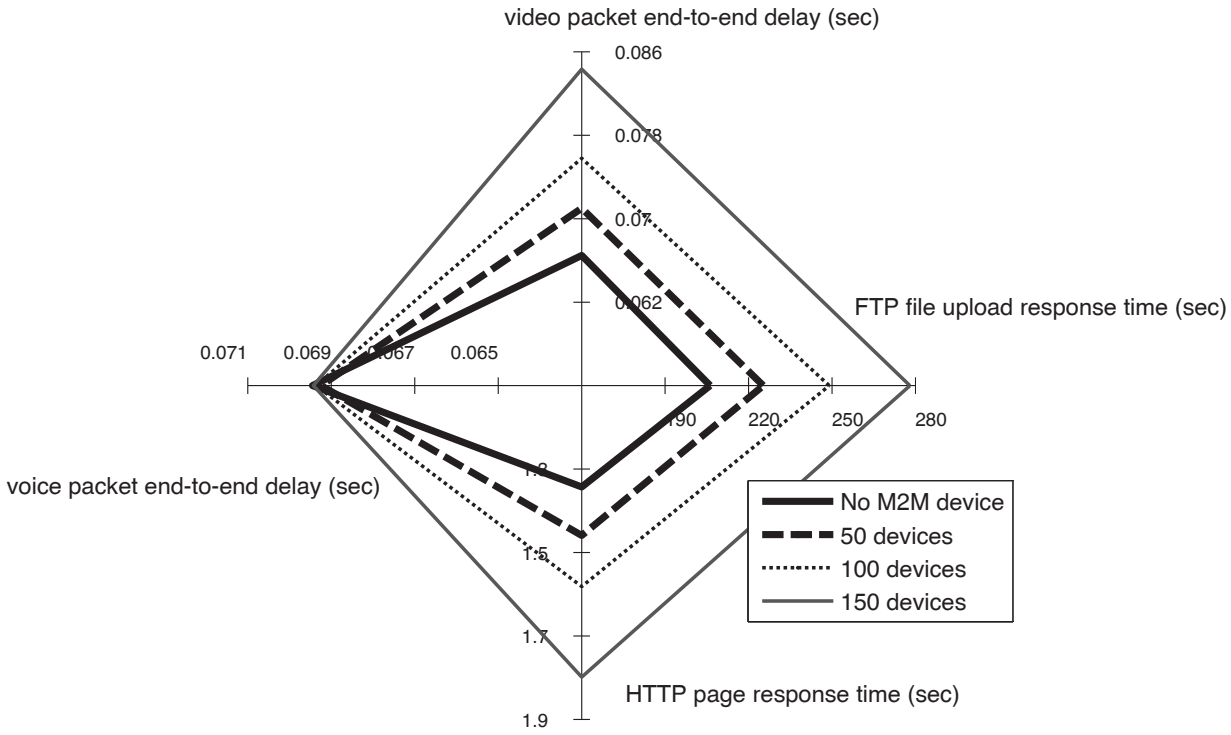

Figure 3.3: Cell QoS performance with e-healthcare data traffic [MWZ+14]

Simulation results in Figure 3.3 indicate that increasing the load of M2M traffic in e-healthcare scenarios has a significant impact on the QoS performance of the regular LTE traffic except VoIP traffic. VoIP users have high priority and increasing the number of M2M devices has no impact on the average packet end-to-end delay performance of VoIP users. The performance of other traffic types degrades with increase in the number of M2M devices. The average packet end-to-end delay of video users, the file upload response time of FTP users and the page upload time of HTTP users increases considerably.

3.7 Conclusion

As anticipated, the M2M data traffic would increase in the coming years and result in high traffic load on the networks. As possible candidates for M2M communication, LTE and LTE-A networks faces huge challenges in terms of capacity and QoS performance. Network operators have to find ways of differentiating between regular traffic and M2M traffic in such a way that the respective QoS requirements are fulfilled for both kinds of traffic. Radio resource allocation mechanisms have to be redesigned, and ways of aggregating traffic before transmitting over the air interface have to be investigated.

4 Relay Node

The 3GPP has been studying several ways of increasing overall data rates and reducing end-to-end latency. In the Release 10 of 3GPP specifications, several new features and functionalities have been introduced for improving the performance of LTE networks. Among major challenges faced by current mobile communication systems like LTE and LTE-A, low throughput for cell-edge users is a significant issue. Several techniques have been investigated to address this problem. The deployment of different low power heterogeneous nodes with the normal micro and macro base stations has been considered as a solution to the issue. Femto cells, pico cells, or Relay Nodes (RNs) can be deployed for this purpose. RNs have been introduced by the 3GPP in Release 10 specifications [33.11]. In this thesis, the focus of investigation is low cost and low power RNs.

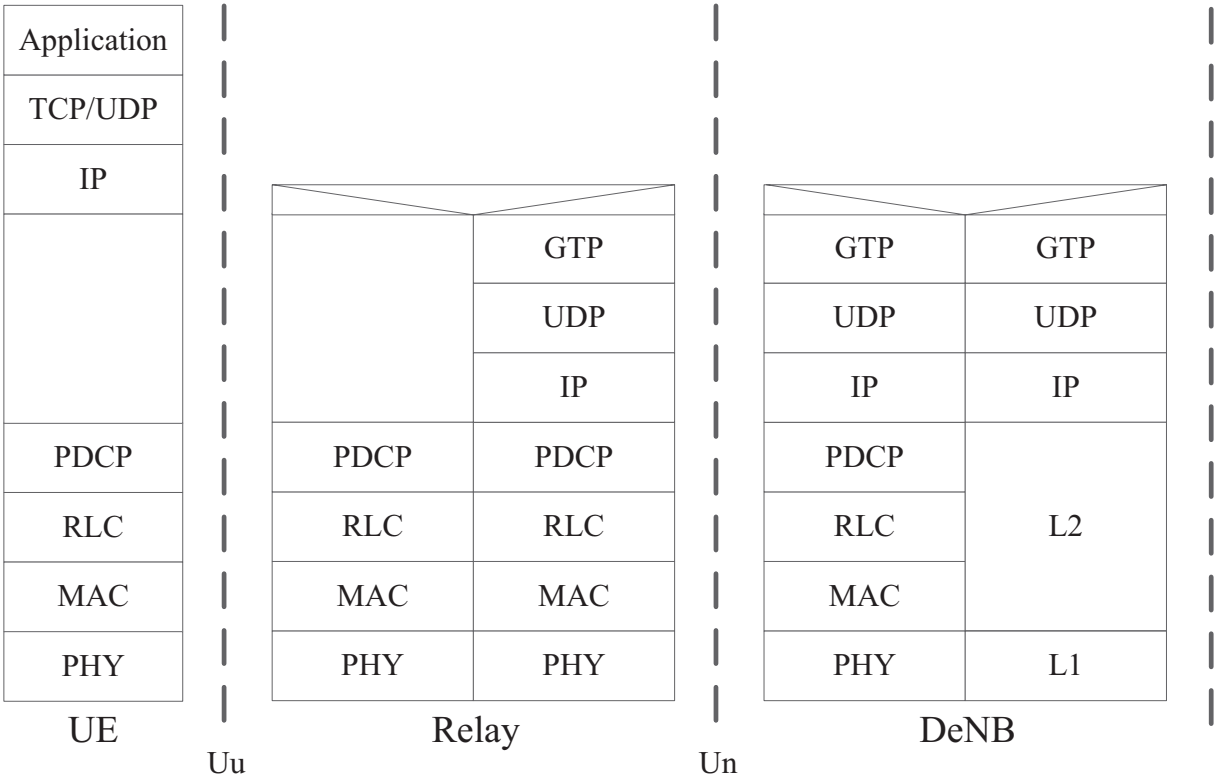

Figure 4.1: Relaying system protocol stack [CWL12]

An RN is a device used for extending the cell coverage area [Cox12]. Relaying is the communication of mobile devices with the network through a node wirelessly connected to an eNodeB over the LTE-A air interface [DPS11]. The RN appears as an eNodeB to the terminal. The RN itself is a low-power base station which is connected wirelessly to the Donor eNodeB (DeNB). The cost of deploying RNs is less than that of femto and pico cells because no additional infrastructure is required for connecting RNs to the DeNB. RNs are designed to improve coverage by placing them at locations in the coverage area having poor channel conditions or coverage holes. The radio scheduling at RNs is done independently of the DeNB scheduling [HT11]. The protocol stack of a relaying system is illustrated in Figure 4.1.

4.1 Relay Node Classification

The classification of RNs can be done based on several criteria such as mobility, technology, resource usage and UE knowledge. Some of the classes of RNs are discussed here.

4.1.1 Mobility Based Classification

Depending on whether an RN is firmly placed at a fixed location or mounted over a moving object like train, vehicle etc., RNs can be classified as either fixed RNs or moving RNs [Bul10].

4.1.1.1 Fixed Relay Node

Fixed RNs are suitable for areas well-known for having coverage issues. In the vicinity on an eNodeB, spots where poor signal quality is regularly experienced can be covered with fixed RNs. Similarly, if coverage holes due to shadowing are identified, especially inside large buildings, fixed RNs can be deployed to provide coverage.

4.1.1.2 Moving Relay Node

In some cases, it is not feasible to deploy fixed RNs as either it might not be possible to provide coverage using fixed RNs or it might be not of economic worth. In cases like emergency, fire, accident or public gatherings, it is imperative that high quality additional radio coverage is required but on temporary basis. Such events occur unexpectedly at random locations, hence fixed RNs are not the answer to

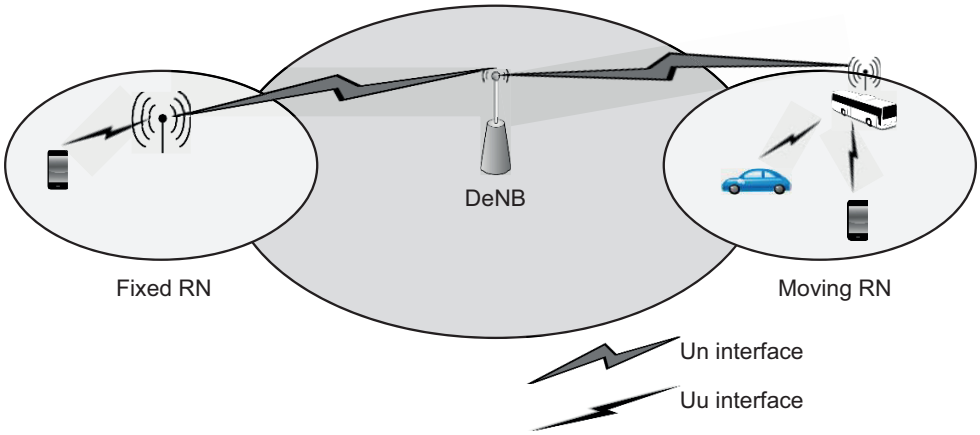

Figure 4.2: Fixed and moving relay

this problem of radio resource requirement. Moving RNs can provide such services, if mounted over police cars, ambulances, fire brigade vehicle etc. Moreover, public transport means can also be used during peak hours of traffic if RNs are mounted on buses or trams and turned on during rush hours. Additionally, passengers travelling in buses, trams, trains or ferries can also be provided connectivity using moving RNs.

4.1.2 Relaying Technology Based Classification

Three classes of RNs have been designed with differences in the adopted technological aspects [ITN10]. These classes vary from each other in terms of how signals are treated after reception and before transmission to next network node.

4.1.2.1 Layer 1 Relay Node

Layer 1 RN is also referred to as a booster or repeater. The role of this relay is to amplify the received signal and then forward it. Only the physical layer functionality of RNs is required for such type of relaying, resulting in low cost deployment as well as short relaying delays. However, layer 1 RN also amplifies the interference from other cells and the noise along with the signal. Thus, the SINR of the amplified signal is reduced and the throughput performance is degraded.

4.1.2.2 Layer 2 Relay Node

The goal of layer 2 RNs is to demodulate and decode the received signal, and then encode and modulate it along with amplification before transmission. In this

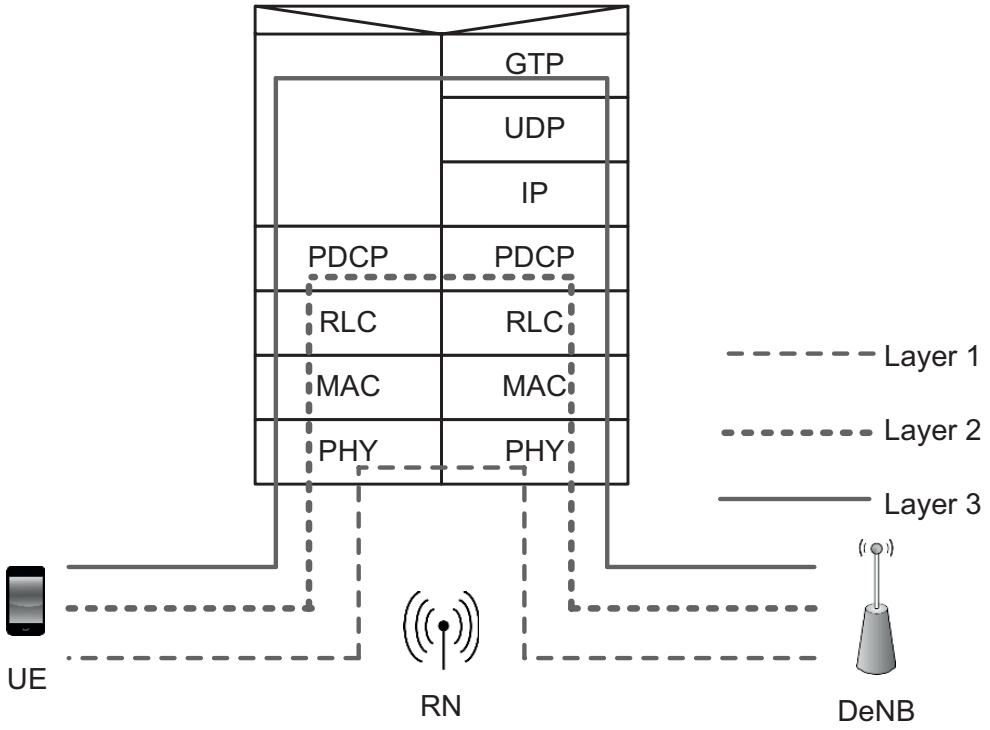

Figure 4.3: Layer 1, 2 and 3 relay

way, the throughput performance can be improved as compared to layer 1 RNs by eliminating the interference and noise. However, additional relaying delay is introduced due to the layer 2 processing.

4.1.2.3 Layer 3 Relay Node

In layer 3 RNs, in addition to the layer 1 and layer 2 processing of the received signal, processes related to layer 3 such as ciphering as well as concatenation, segmentation and reassembly of data are also performed, which are eNodeB functionalities. This results in even better throughput performance as compared to layer 2. However, layer 3 RN provide higher throughput at the cost of increased relaying delays. The 3GPP has decided to finalize the standardization of RN with layer 3 as the relaying technology.

4.1.3 Air Interface Based Classification

In uplink transmissions, RNs transmit to the DeNB on the backhaul link and receive from UEs on the access link. Both these transmissions have to be achieved by availing the same frequency bands. Similarly, in downlink case, signals on the backhaul link from the DeNB to RNs and on the access link from RNs to UEs

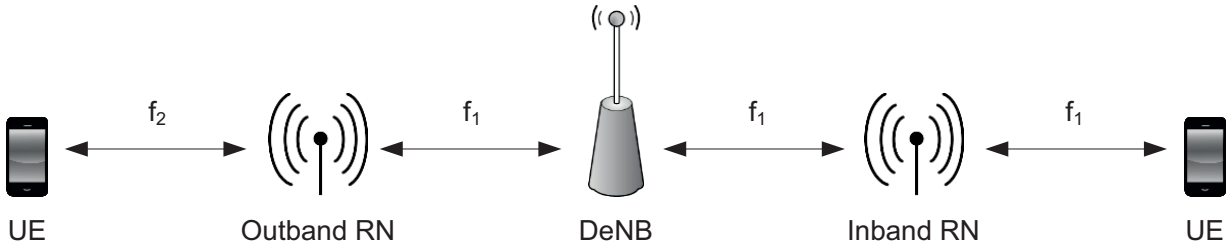

Figure 4.4: Outband and inband relay

have to share the same frequency band. So there are two possibilities of sharing the spectrum, either dividing the frequency band with outband operation, or sharing the time between the two links with inband operations. According to 3GPP specifications, both inband and outband operations of RNs are permitted.

4.1.3.1 Outband Relay Node

In outband operations, both the access link over Uu interface and the backhaul link over Un interface operate on different carrier frequencies. Outband operations may not always be feasible to adopt for RNs since the two links require separate frequency carriers which are scarce resources.

4.1.3.2 Inband Relay Node

The inband operation indicates that the backhaul link as well as the access link are using the same carrier frequency. Inband operations are possible if the two links are time multiplexed. If the separation between the transmit and receive antennas at RNs to avoid interference can be achieved, then both the links can use the whole frequency without time multiplexing [HKB11]. For example, in case of indoor RNs, it is possible to deploy the backhaul link antenna outside and the access link antenna inside the building.

4.2 Solutions for M2M Communication in LTE-A

As described, the primary reason for the design of RN architecture in LTE-A is to increase the cell edge throughput. The narrowband M2M applications generate data traffic having small sized packets. The M2M devices for such applications are generally designed to send very few bits after a certain number of seconds or even hours. However, the 3GPP has standardized LTE-A in such a way that radio resources are allocated to users with a PRB as the smallest resource unit. A PRB

is made up of 12 OFDM sub-carriers in the frequency domain. A PRB can be utilized to transmit several hundred bits of data in favorable channel conditions. The radio spectrum is a scarce resource. The network operators are able to obtain licenses for spectrum by paying huge amount of money to start mobile services. Therefore, allocating an entire PRB to a single M2M device can result in severe reduction in the overall spectrum efficiency and resource utilization. In the light of this discussion, it is quite perceptible that modifications in the design of the LTE-A network architecture and functionalities is required in order to be able to cope with the growing M2M traffic with QoS assurance. Before discussing the role of RNs for M2M communication, some possible solutions for M2M traffic integration into LTE-A networks are presented.

To deal with the problem of serving M2M traffic in LTE-A network without compromising spectral efficiency of the transmission, a number of possible solutions can be investigated for M2M communication. Some of these solutions are as follows [MZC+14].

- Each M2M devices can be designed to delay the transmission and aggregate a sizeable amount of data together in its transmission buffer before sending it to the LTE network. This would clearly increase the efficiency of the resource usage, but the M2M devices are expected to send data after several seconds or even hours and such a long delay in transmitting the data for achieving aggregation of several data packets together might have a negative impact on the service provision of the M2M device. In terms of practical operations, this solution may not always be feasible.

- Looking at the technical possibilities, the LTE network can be redesigned to allocate the M2M devices with less than a single PRB. Thus, 1 sub-carrier can be allocated to a device. But the drawback of this redesign is that the 3GPP standards have to be modified for implementing this change. Moreover, reducing the resource allocation unit size to a subcarrier would also require that more OFDM gap intervals are inserted, otherwise the inter symbol interference would be unavoidable. The greater gap intervals would lead to a reduced spectrum capability. Due to these reasons, this solution can also not be considered to be practically feasible.

- M2M traffic emerging from a large number of M2M devices can be aggregated and multiplexed together for the transmission over a single PRB, or if required many PRBs. In this way, the disadvantages of the above mentioned solutions can be averted. In order to achieve this, the technical challenges related to this solution have to be addressed before implementing it. The

most obvious challenge is to determine how to aggregate the M2M traffic before transmission towards the eNodeB. Additionally, it has to be formulated that how the traffic should be multiplexed and demultiplexed. The role of various LTE-A protocols has to be defined for implementing the scheme.

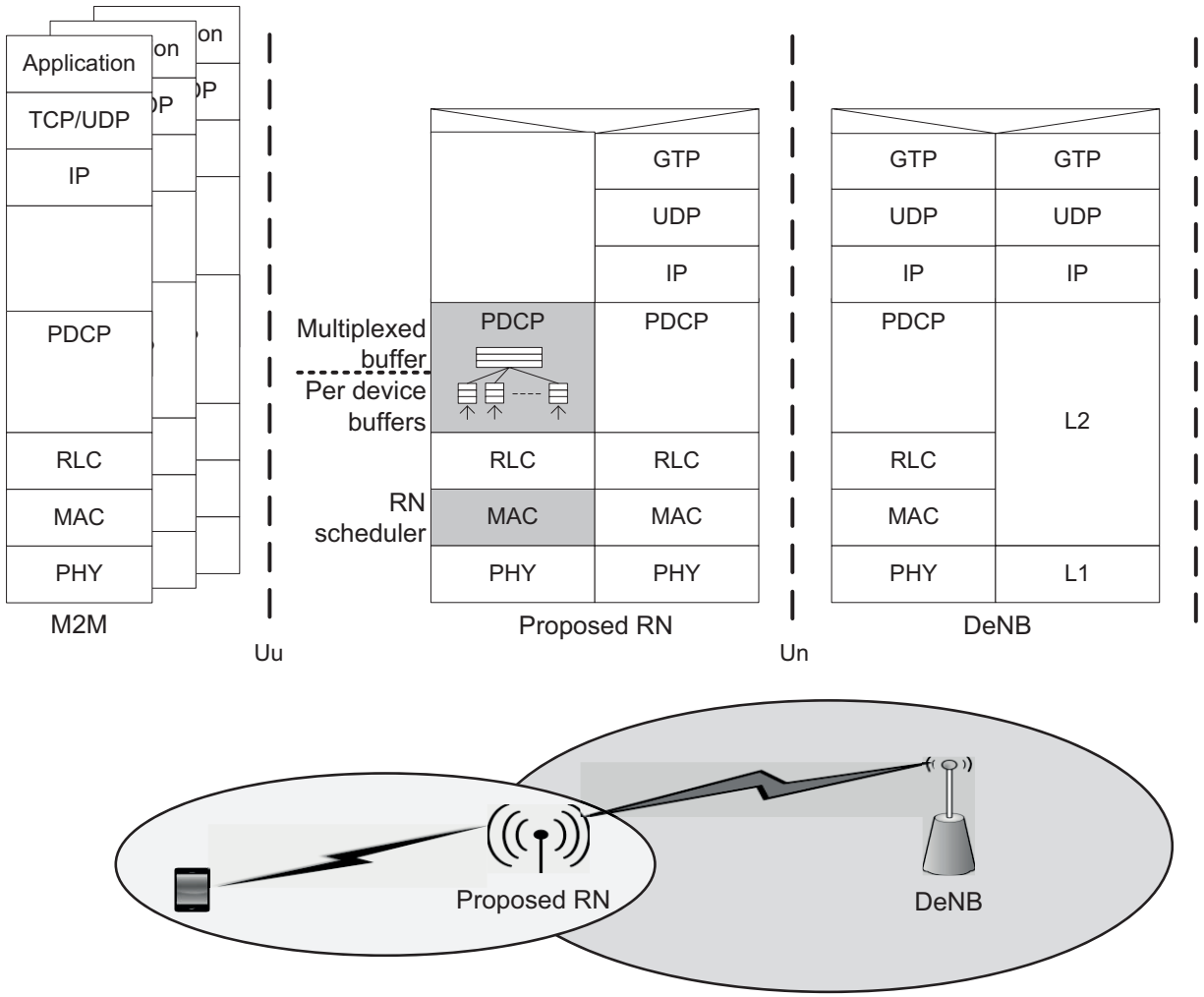

Figure 4.5: Protocol stack for RN based data multiplexing framework [MZC+14]

4.3 Relay Node for M2M Communication

In this thesis, the wireless inband layer 3 fixed type of RNs are considered as the focus of research. This type of RNs are proposed as the solution for aggregating and multiplexing the data from a large number of M2M devices. The RNs can be deployed with low implementation costs because additional wired backhaul links are not required. The RN uses the wireless spectrum that is provided by

the operator to the DeNB. The RN functionality can be modified to aggregate and multiplex the M2M traffic before transmission to the DeNB.

To achieve efficient integration of M2M uplink traffic into the LTE-A network, some modified functionalities are required to be implemented in the wireless in-band layer 3 fixed RNs. A summary of these functionalities is given here.

- The *RN scheduler* is required for scheduling the radio resources over the RN Uu interface. This scheduler should allocate resources to the M2M devices of various M2M traffic classes. To achieve efficient M2M transmission, the scheduler should be aware of the contrasting M2M QoS characteristics of the devices and correlate access link scheduling with the varying channel conditions of the wireless backhaul of the Un interface between RNs and the DeNB.
- The *Multiplexing algorithm* in the RN is needed for aggregating the M2M traffic from different devices and multiplexing them before transmission to the DeNB. The algorithm has to work hand-in-hand with the RN scheduler so that the data is multiplexed according to the available radio resources for RNs.

Thus, instead of requesting radio resources from the DeNB for the individual M2M devices, the resources are requested by RNs for a group of M2M devices with data multiplexed at the PDCP of RNs. This can be achieved if the aggregated IP packets from several M2M devices are multiplexed into a single large PDCP packet. At the DeNB, the resource request for this large multiplexed packet would be seen as a single radio resource request. Hence, it would be possible to allocate one PRB (or if required, more PRBs) to several M2M devices, resulting in the enhancement of spectral efficiency many folds.

4.4 OPNET Simulation Environment

The simulation environment explained in section 2.1.6 is enhanced in such a way that instead of a single eNodeB in the environment, a seven cell environment is developed using the OPNET Modeler (Figure 4.6) with a frequency reuse factor of 1. The enhanced model is based on improved mobility and channel models as well as revised scheduling algorithms. The new model not only emulates the effects of path loss, slow fading, fading and noise, but also the interference from the neighboring eNodeB caused by the mobile devices in the uplink and eNodeBs in the downlink transmissions. The fading effects are modelled by using the path loss with slow fading maps and the fast fading model for macro cells of the open

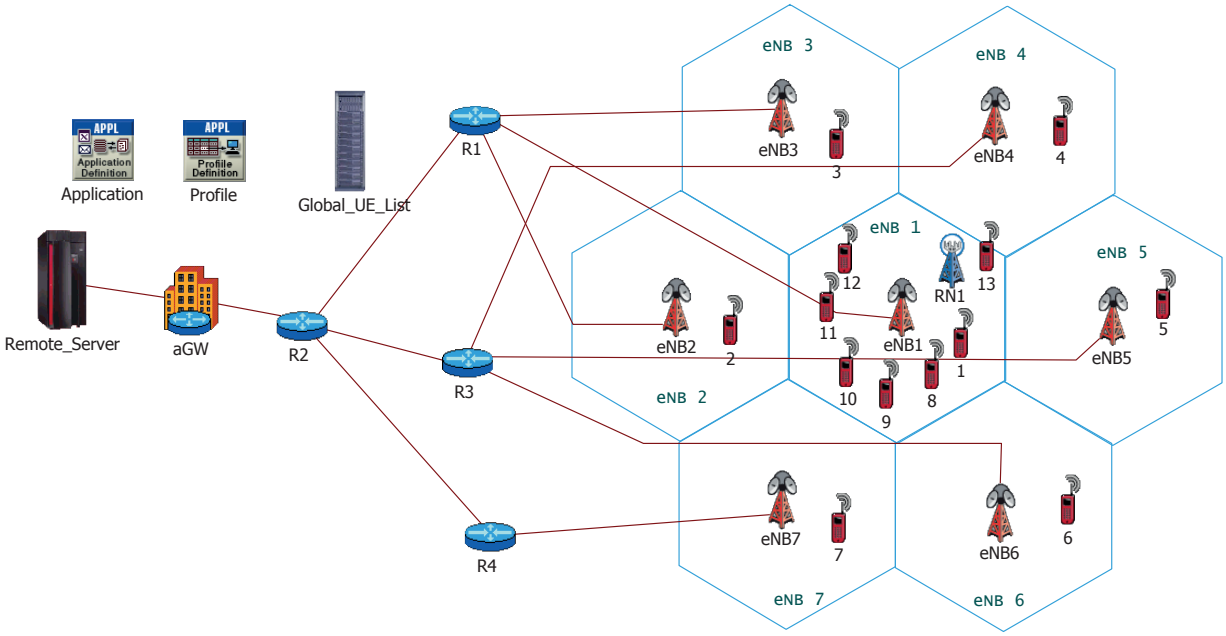

Figure 4.6: OPNET LTE-A simulation model

source Vienna Simulator [IWR10]. Due to the frequency reuse factor of 1, if a UE transmits uplink signals by using certain PRBs, this signal would cause interference to the transmission of the UEs of the neighboring eNodeBs using these PRBs. The interference also depends on the location of the interfering terminal. Cell edge UEs would cause greater interference to the neighboring cell. The scheduling algorithms should also be designed in such a way that the signal interference is also considered for resource allocation.

4.5 Relay Node Implementation

The RN node model is developed with two air interfaces, the Un and Uu. The Uu and Un protocol stacks are developed in accordance with the LTE-A protocols and end-user protocols. The Uu protocol stack consists of PDCP, RLC, MAC and Physical PHY layers (labelled in Figure 4.7 as pcdp_0, rlc_0, mac_0 and phy_0 respectively). The RN scheduler is located at the MAC layer of the Uu interface where resources are allocated to the users served by the RN.

The proposed scheme for packet data aggregation and multiplexing into large packets is implemented at the PDCP layer of the Uu interface. At the Un side, the protocols are the GPRS[1] Tunneling Protocol (GTP), UDP/IP, PDCP, RLC, MAC and PHY layers. The GTP protocol is implemented in such a way that the

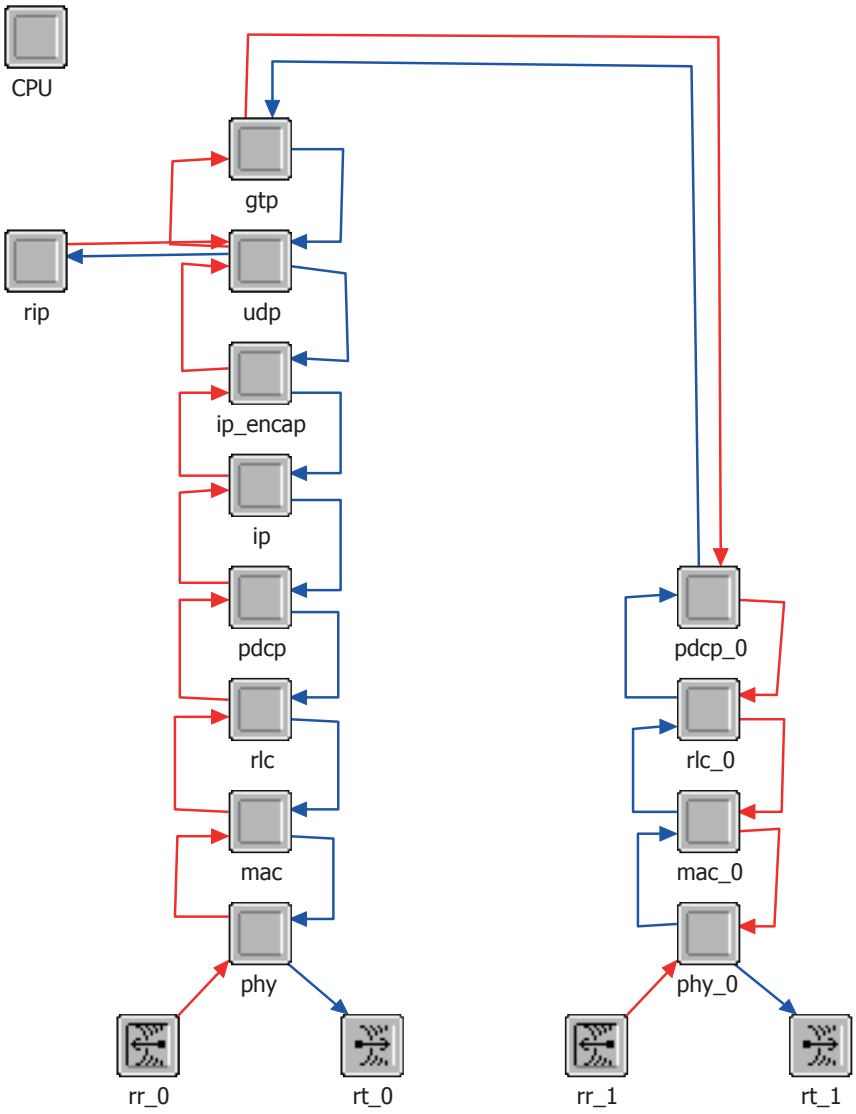

Figure 4.7: RN node model

data packets are tunneled at the RN and detunneled again at the DeNB. Similarly, downlink packets are tunneled at the DeNB and detunneled at the RN.

The details of the implemented RN MAC scheduler and the PDCP aggregation and multiplexing scheme are provided next.

4.5.1 DeNodeB Scheduling with Relay Node

The literature survey of the topic of radio resource allocation with RNs in the coverage area resource allocation reveals the novelty of the topic. Few authors have already proposed scheduling schemes for the DeNB MAC layer, where resources

[1]GPRS stands for General Packet Radio Service

are allocated to both normal UEs and RNs [CG14], [GRG13], [dMSBS13]. The previous work done on scheduling schemes assumes inband and outband relaying for layer 3 RNs. In this thesis, the primary issue is to design fixed RNs for M2M data traffic multiplexing. Thus, the case of M2M devices served by RNs with backhaul link and access link antennas having good separation is considered. For example, the access link antenna can be inside a warehouse building with sensors deployed for monitoring the temperature of several points and the backhaul link deployed outside the building. As a result, the RN inband operation can be performed without time division of the resources.

An LTE-A uplink scheduling scheme, Channel and QoS Aware (CQA), was described in section 2.2. This scheduling scheme is now modified to consider RNs as well along with other normal UEs while distributing the PRBs. The uplink scheduler in section 2.2 is time and frequency domain decoupled. In time domain, the candidate lists for GBR and non-GBR bearers are created. Both lists are sorted according to the TDPS metric values of the bearers so that the bearers having high metric values are placed at the top of the list for high priority in scheduling. The $W_{i,a}(t)$ parameter indicating the QoS weight of bearer a of user i at time t of the Weighted Proportional Fair (W-PF) algorithm given in (2.10) is modified to (4.1).

$$W_{i,a}(t) = \frac{R_{\min,a}}{R_{\text{avg},i,a}(t)} \tag{4.1}$$

In (4.1), the parameters related to bearer delay are eliminated from the equation due to the fact that in uplink data packet transmission, the information about packet delay at the PDCP buffer of the UE (or RN) is not available at the eNodeB. The Buffer Status Report (BSR) provides information only about the length of the buffer at the eNodeB. Although, the length of the buffer can be used to estimate the delay [BMS12], but since the scheduling scheme in the CQA scheduler is already designed to prioritize the delay sensitive GBR bearers by creating a separate GBR bearer list, so the delay estimation is not critical. Hence the delay estimation is not considered for the scheduling decisions.

In the TDPS, RNs are considered as users with uplink data buffered for transmission. In order to make sure that RNs gets PRBs in each TTI (assuming that RNs are serving a huge number of M2M devices and data is always available for transmission), RNs are considered as GBR users. To assure fairness to other bearers, especially the non-GBR bearers, the maximum number of PRBs attained by each RN $N_{\max}$ is restricted.

4.5.2 Relay Node Scheduling

The RN scheduler located at the MAC layer of RN Uu interface is designed to allocate radio resources to the UEs and M2M devices connected to the RN through the access link. RN has been standardized by 3GPP and several eNodeB functionalities are performed by the RN, but the primary difference between an RN and an eNodeB is that the RN is connected to the network over the backhaul link wirelessly. On the other hand, the eNodeB is connected to the core network with a wired interface. As explained, time division or frequency division multiplexing are the two options for avoiding interference between the access and backhaul links.

In case of M2M traffic, data transmitted by devices is in the form of small sized packets. The packets are transmitted with intertransmission time in the order of seconds. However, the number of M2M devices is assumed to be huge. Therefore, the arrival of packets is expected to be regular at the RN due to the huge number of devices. In contrast to regular traffic like FTP, where a single UE would be sending bursts of packets, the M2M traffic would have bursts with packets from different devices. The round robin scheduling approach is adopted for such type of M2M traffic. All the active devices ready to transmit packets are scheduled in a round robin fashion.

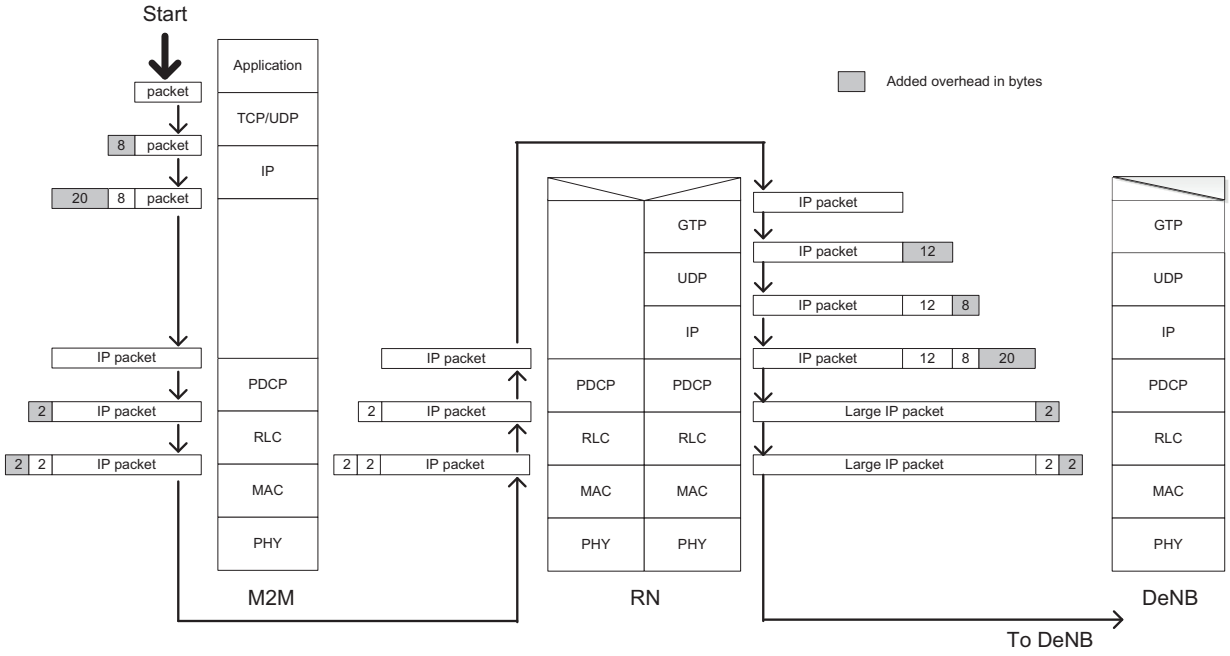

Figure 4.8: Packet flow from M2M device to DeNB with relaying

4.5.3 Relay Node Aggregation and Multiplexing Scheme

The aggregation of packets from several devices at the RN PDCP layer on the Uu side is performed for multiplexing several small sized IP packets into one large IP packet. The goal of multiplexing is to utilize the radio resources of the Un interface between the RN and the DeNB efficiently for M2M traffic. The reason for choosing the PDCP layer for data multiplexing is to get rid of the PDCP, RLC, MAY and PHY overheads (if any) on the Uu side before multiplexing. Otherwise, multiplexing could also be possible at the other Uu layers. The flow of a packet from the M2M device over the RN to the DeNB is illustrated in Figure 4.8. The packet originates at the application (APP) layer of the M2M protocol stack.

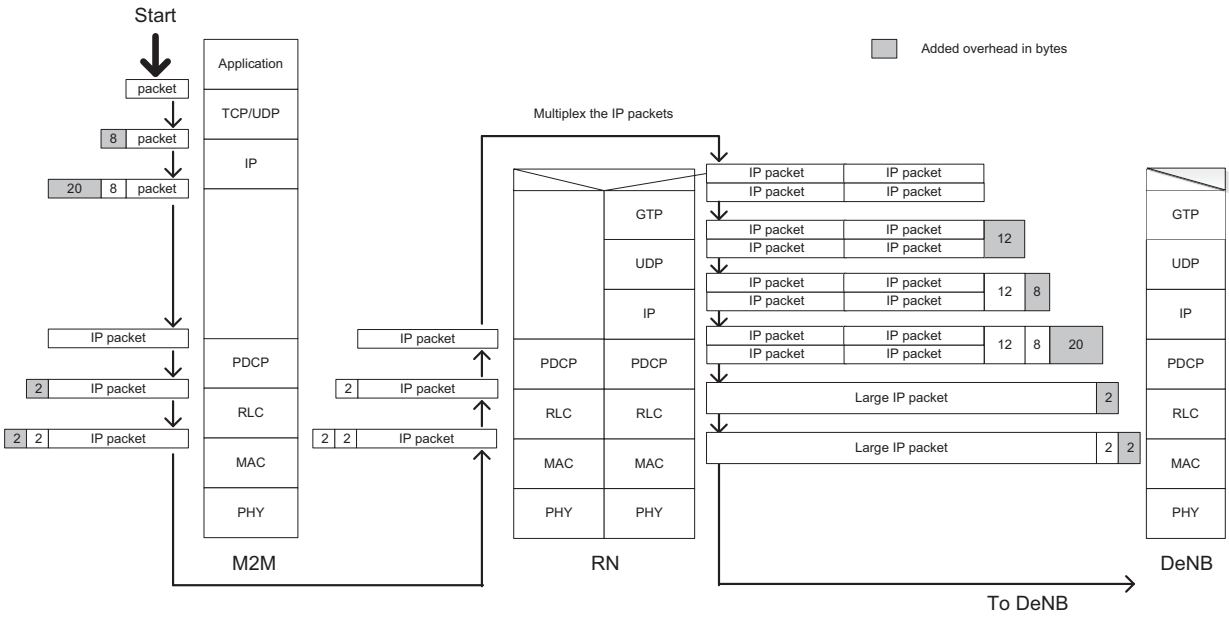

Figure 4.9: Packet flow from M2M device to DeNB with multiplexing at the RN

At the PDCP layer of the RN Uu, the IP packets from various devices are multiplexed by initially placing them in an aggregation buffer as depicted in Figure 4.9. For packets of several QoS classes, individual buffers can also be created if the DeNB scheduler is capable of differentiating between the traffic classes. For the time being, it is assumed that a single buffer is used for aggregating the packets for multiplexing. Two major performance parameters are considered while multiplexing the data, i.e., throughput and QoS. In order to maximize the throughput, the packets are aggregated until sufficient packets arrive at the RN Uu PDCP layer such that the arrived packets can fully utilize the available TBS provided by the DeNB. In this case, the DeNB cell throughput performance would improve because the RN would be able to use the available resources with maximum uti-

lization. In other words, the buffer size reaches a certain value $n_{\max}$. This value should be the TBS achievable under the channel conditions of the RN and the maximum number of PRBs $N_{\max}$ that the DeNB scheduler allocates to the RN. But the drawback of throughput maximization is that the delay sensitive traffic might not achieve the desired end-to-end delay performance. In that case, a mechanism is required to make sure that packets are not buffered for such a long duration that the QoS performance is degraded [DRVV11]. For this reason, a timer is introduced at the buffer. The role of the timer is to wait for the packet arrivals for a certain time duration $T_{\max}$, and start the multiplexing process at the expiry of this time duration. In this way, the multiplexing would be performed during a certain time duration $T_{\max}$, after which the packets are multiplexed into large IP packets. The multiplexing and timer mechanism is graphically illustrated in Figure 4.10.

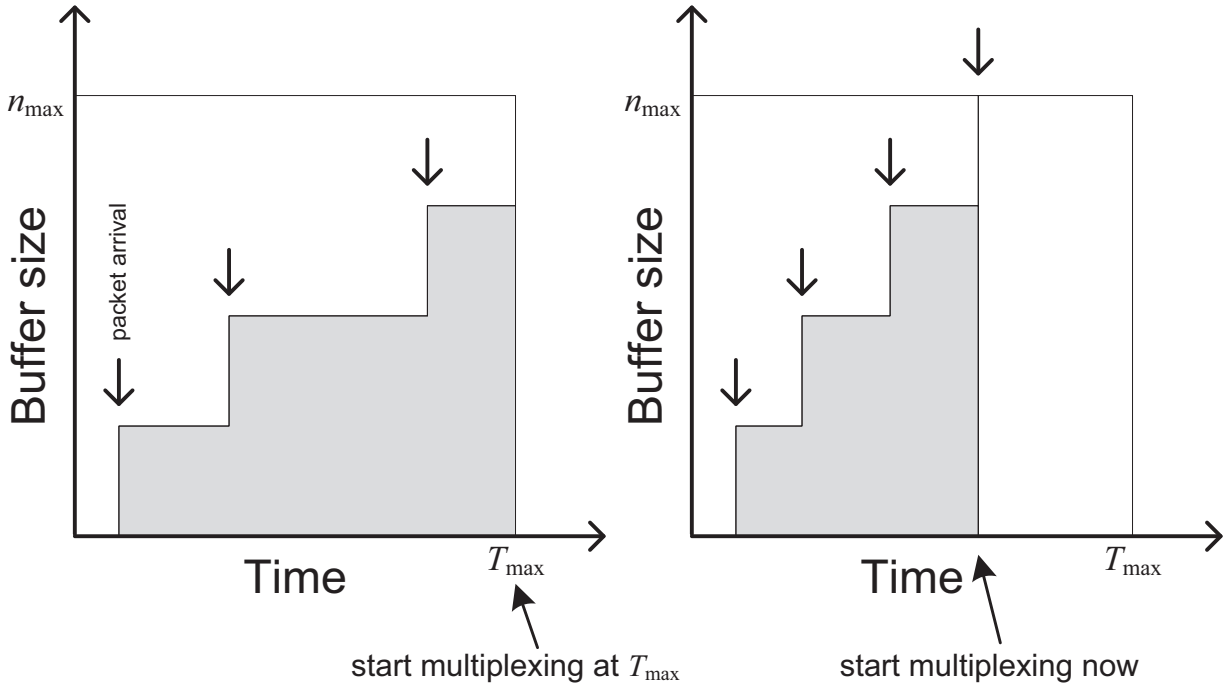

Figure 4.10: Multiplexing due to timer expiry and maximum buffer size

4.6 Simulation Parameters, Traffic Models and Results

The settings for various simulation parameters and the traffic models for regular LTE traffic as well as the M2M traffic are listed in Table 4.1. As described, the simulation environment is modelled according to the LTE-A protocols and functionalities. The eNodeBs with omni-directional antennas are used for simulation

purposes. The model consists of a hexagonal regular grid setup where the center cell is surrounded by another tier of six cells in the simulation environment (Figure 4.6). The RN is connected to the central eNodeB labelled as eNB1 in Figure 4.6. Since the RN is fixed, therefore the channel conditions for RN over the Un interface would remain consistent throughout the simulation duration.

Parameter	**Setting**
Cell layout	7 eNodeBs with single cell
System bandwidth	25 PRBs ($\approx$ 5 MHz)
Inter eNodeB distance	500 m
Max uplink transmit power	23 dBm
Channel models	Vienna simulator [IWR10]
Noise per PRB	-120.447 dBm
Noise floor	9 dB
Power Control	FPC, $\alpha = 0.6, P_0 = -58$ dBm
Timer expiry $T_{\max}$	9 msec
VoIP traffic model	
Silence/talk spurt length	Exponential(3) sec
Encoder scheme	GSM EFR
Video traffic model	
Frame size	Constant(1200) byte
Frame inter-arrival time	75 msec
FTP traffic model	
File size	Constant(20) Mbyte
File inter-request time	Transmission completion of last file
M2M traffic model	
Message size	Constant(29) bytes (at PDCP)
Message inter-transmission time	1 sec

Table 4.1: Main simulation parameters and traffic models

The simulation results are presented in a categorized manner. The simulation results categories include coverage enhancement with RN, M2M service without RN, M2M service with RN and multiplexing, impact of M2M traffic on LTE-A regular traffic, and impact of timer.

4.6.1 Coverage Enhancement with Relay Node

One of the primary goals of designing RNs is to extend cell coverage and improve cell edge throughput and QoS performance. The UEs at the cell edge of an eNodeB usually experience poor channel conditions. The uplink Power Con-

Number of PRBs	1	2	3	4	5
MCS 16	41	79	121	161	201
MCS 18	47	97	145	193	249
MCS 20	55	113	173	233	293
MCS 22	65	133	201	269	333
MCS 24	73	149	225	301	373
MCS 26	89	185	277	373	469

Table 4.2: TBS in bytes for several MCS [36.10a]

trol (PC) plays an important role in determining the maximum number of PRBs a UE can be allocated by the scheduler. The cell edge users are generally not able to utilize the available bandwidth because PC strives for reducing the UE battery consumption and interference towards the neighboring cells. This results in bad cell edge throughput and QoS performance. RNs can play a vital role in improving the performance of cell edge UEs.

Scenario parameter	Setting
RN location MCS	16
UE location MCS	4
Maximum PRBs for RN	25 PRBs
Video UEs	10, 20, 30, 40, 50, 60

Table 4.3: Scenario specific simulation parameters

The simulation scenarios are divided into 2 categories. The overall cell performance of cell edge users, directly connected to the eNodeB, is evaluated in the 1st category of scenarios. In the 2nd category, the performance of user connected via RN is evaluated. The UEs and the RN are stationary, which means that the channel conditions are consistent throughout the simulations because of the fixed positions of the nodes. The location of the RN corresponds to MCS 16 (see Table 4.2 [36.10a]), while the UEs location results in an MCS of 4 if directly connected to the eNodeB. In all the subscenarios of the RN based eNodeB access category, the RN is allowed to utilize all the available PRBs in cell, since no other users are deployed in the cell. The scenario specific simulation parameters are given in Table 4.3.

Both categories, including direct eNodeB access and via RN access, are further subdivided into subscenarios with only *video* UEs in the cell. In the 1st subscenario, only 10 UEs with video uplink traffic are deployed. The number of users

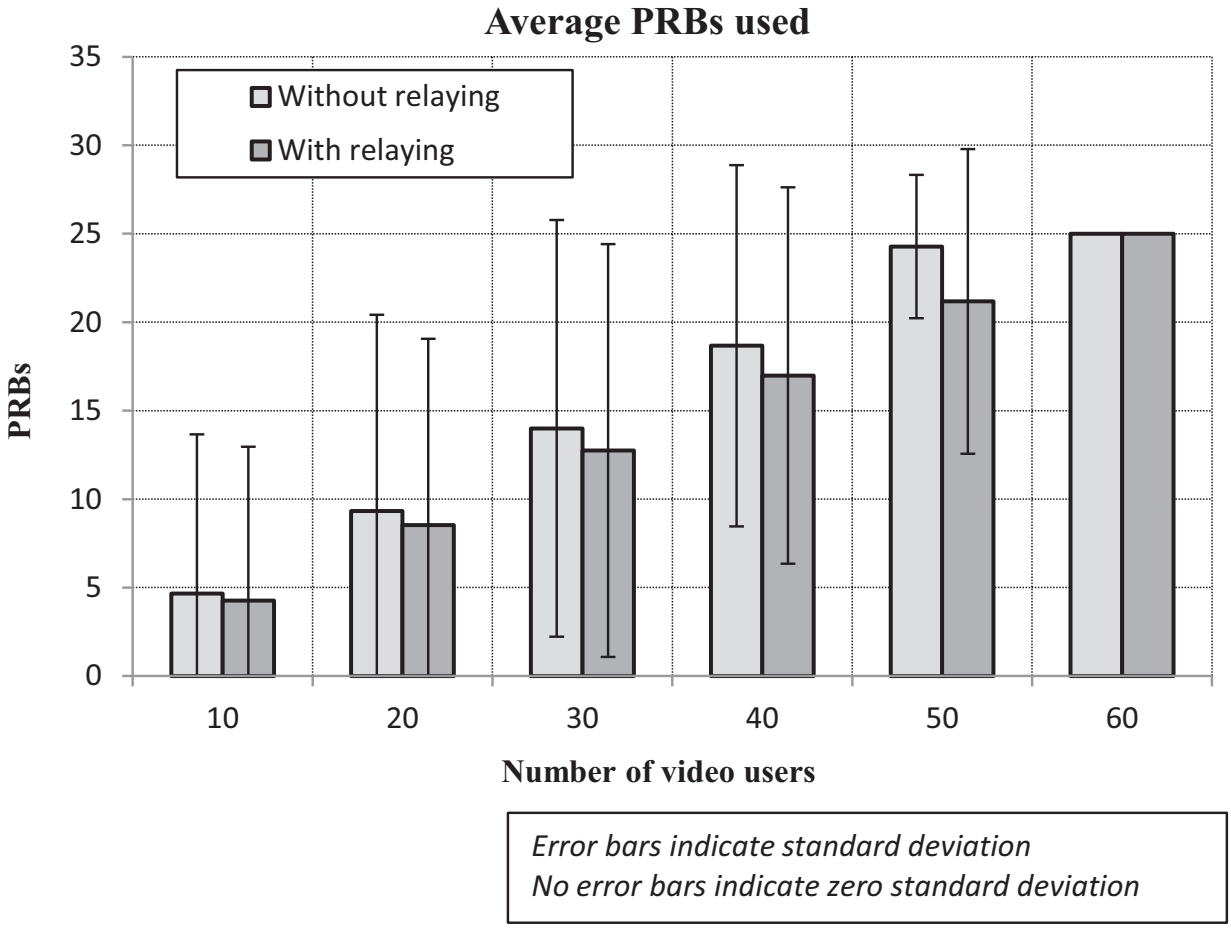

Figure 4.11: Average Uu PRB usage with and without RN for video traffic

is incremented by 10 in each subsequent scenario. The cell edge coverage improvement for video UEs deployed at the cell edge with poor channel conditions is illustrated in Figure 4.11 to Figure 4.13. The error bars in the graphs represent the standard deviation of the results. The 95% confidence intervals for all results are provided in Appendix.

Figure 4.11 depicts the average PRBs used by the video users in different subscenarios of both the aforementioned categories. A comparison of PRB usage with and without RN shows that an improvement of 9 to 15 percent is achieved when the RN is deployed for the cell edge users. The PRB usage for low load subscenarios, i.e., 10, 20, 30 and 40 video users at the cell edge is much lower than 25 PRBs, which is the total bandwidth. In the case of 50 users, the average PRBs requirement in the direct eNodeB access category is quite close to the maximum. This implies that it is highly probable that the PRBs requirement might have gone above 25 during the simulation. In that situation, all the data ready for transmission cannot be served immediately by the eNodeB. Some of the data has to wait at

least until the next TTI. This results in higher packet end-to-end delay times when the load is high. For 60 users, the minimum requirement is 25 PRBs during the whole simulation duration which is the whole system bandwidth. The standard deviation in the graphs decreases with increasing traffic load. With 60 users, the load is beyond the maximum capacity and hence the standard deviation reduce to zero since 25 PRBs are required in every TTI.

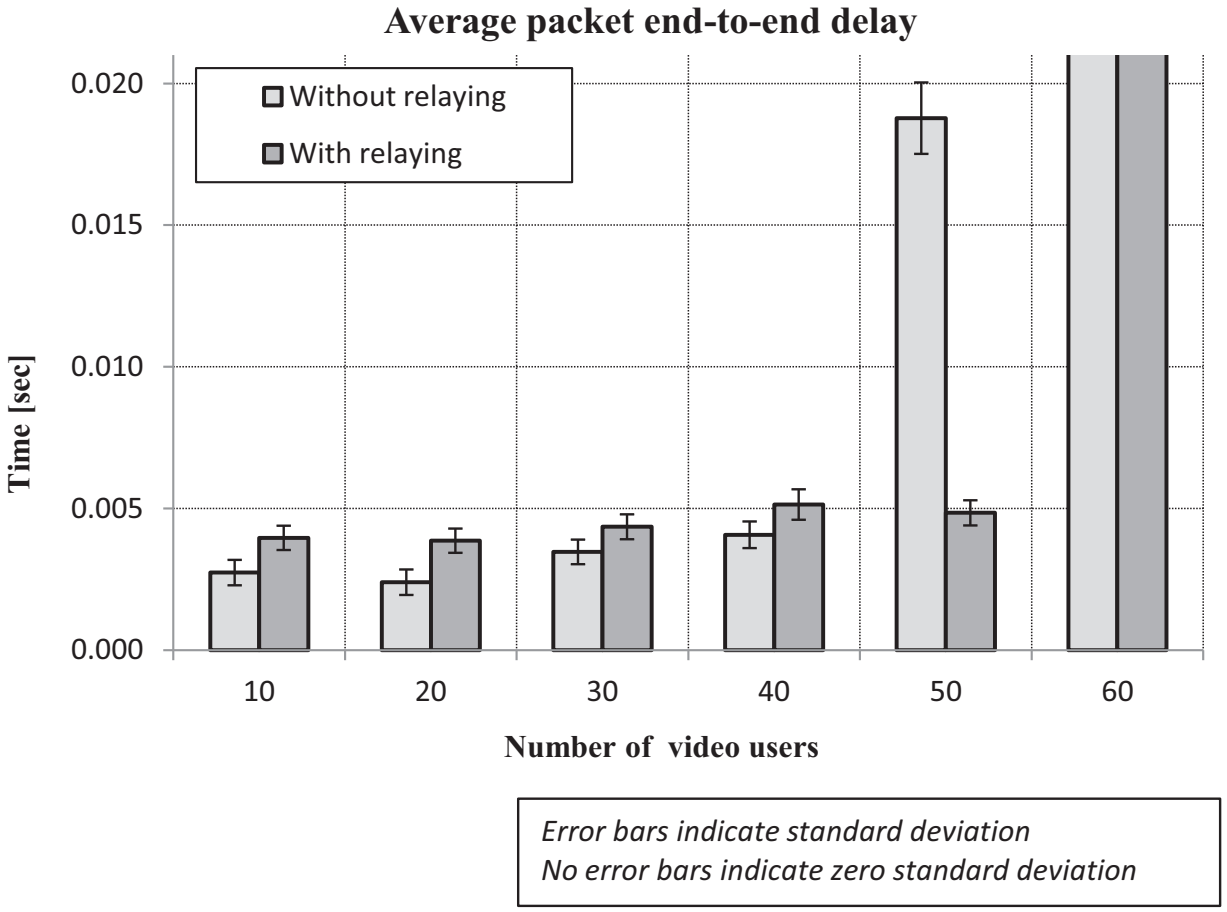

Figure 4.12: Average packet end-to-end delay with and without RN for video traffic

In Figure 4.12, the average uplink packet end-to–end delay of video users is graphically shown. It can be observed that in the low load subscenarios, the packet end-to-end delay resulting for both categories of simulations are similar. The scenarios of access via RN have a slightly higher end-to-end delay due to the additional delay caused by relaying. In the subscenario of 50 UEs, the result for direct access case deteriorates as the system capacity has almost been reached, which can be seen in Figure 4.11 showing that almost all the available PRBs are utilized. However, the end-to-end delay result for the category of access via RN for 50 UEs is still good, despite the higher load because the PRBs are now capable of transmit-

ting more bits due to the proximity of the RN resulting in higher TBS. Finally, in the subscenario of 60 UEs, the results for both categories deteriorate significantly and are beyond the scale of Figure 4.12. The standard deviation in the delay results is small when the traffic load is low. This is due to the fact that free resources are available most of the time without any queuing delays. However, when the load increases, the impact of queuing delay also plays a role in the end-to-end delay of packets. Thus the standard deviation increases for high load subscenarios.

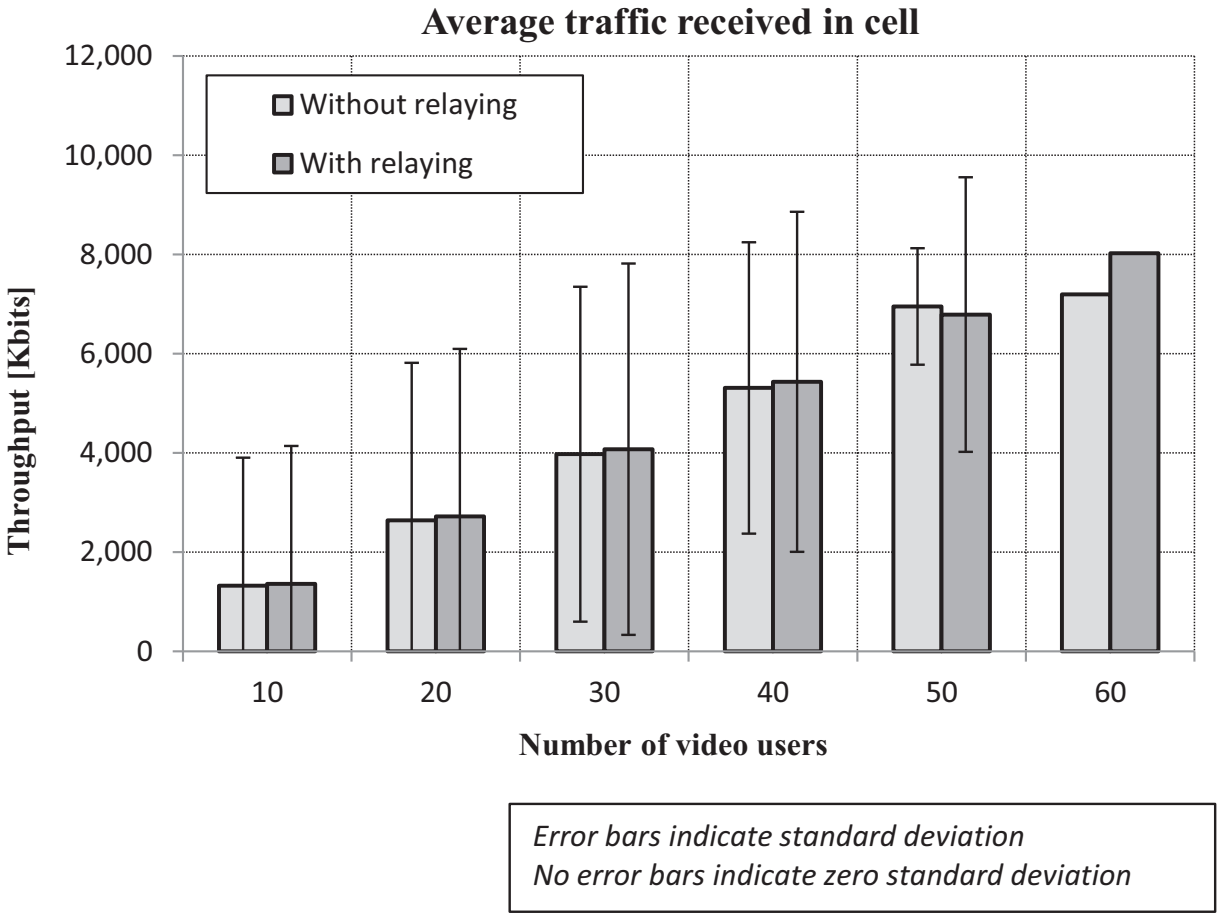

Figure 4.13: Average video traffic received in cell with and without RN

The average data traffic received in the cell for all the subscenarios is depicted in Figure 4.13. Increasing the load results in larger received cell traffic until the maximum load capacity is exceeded. When the number of UEs is increased to 60, the average received traffic is already at the maximum level as the the maximum capacity has been exceeded. An unusual observation in subscenarios with 50 users is that the average received cell traffic is higher when the users access the eNodeB directly, as compared to the access via the RN. The reason behind this peculiar outcome is that 50 users with direct eNodeB access require 25 or more PRBs

frequently during the simulation and are unable to transmit all the data ready in the buffer due to the small achievable TBS because of poor channel conditions. As a result, in a TTI when more than 25 PRBs are required, a part of data is transmitted and the remaining data wait until the next TTI to get resources from the scheduler. Thus, the average PRB usage increases for this subscenario significantly. In case of RN, the need of 25 or more PRBs arises less frequently due to better channel conditions of RN. Therefore, if data are successfully transmitted in a TTI, then it is also possible that in the next TTI, very little resources are requested for newly arriving packets in the buffer. Thus, fewer PRBs are required for the transmission of same amount of data. Hence, for direct eNodeB access, additional overhead is required to be sent along with the data because of the increased PRB usage and therefore, the air interface traffic appears to be higher in this particular subscenario because the additional overhead is also considered as a part of the air interface traffic. The standard deviation in the received traffic results is higher for low load scenarios. As load is increased and the maximum capacity is approached, the standard deviation reduces. When the load is beyond the maximum cell capacity, the standard deviation ceases to exist.

The simulation setup presented for video users is used for the performance evaluation of FTP users in terms of file upload time performance. Video users are replaced by FTP users and there is only FTP uplink traffic in the cell. The two categories of scenarios, eNodeB access with and without RN, are subdivided into subscenarios with different cell loads. The number of FTP UEs in the 1^{st} scenario is 5 UEs at the cell edge and the location corresponds to an MCS of 4. In each subsequent subscenario, the number of FTP users is increased by 5. The RN location in the 2^{nd} category of scenarios is in accordance with an MCS of 16. The results for the average file upload time for uplink FTP transmissions are illustrated in Figure 4.14. The result clearly show that deploying an RN helps in reducing the file upload time for the cell edge FTP UEs significantly. The improvement in file upload times ranges from approximately 9.8 percent to 13.8 percent in the given scenarios. The error bars depict the standard deviation of the file upload time results for the two categories of scenarios. The standard deviation increases when the traffic load in increased. However, as the maximum cell capacity is reached, the standard deviation reduces again for the subscenario of 25 users without relaying. With relaying, since the cell capacity is enhanced and the cell load is below the maximum capacity for 25 users, the standard deviation is still quite large.

Based on the discussion of the results, one can clearly observe that the role of the RN is quite imminent in LTE-A networks for coverage extension and cell edge performance. Dead spots can also be served quite effectively with the help of RNs in a similar manner.

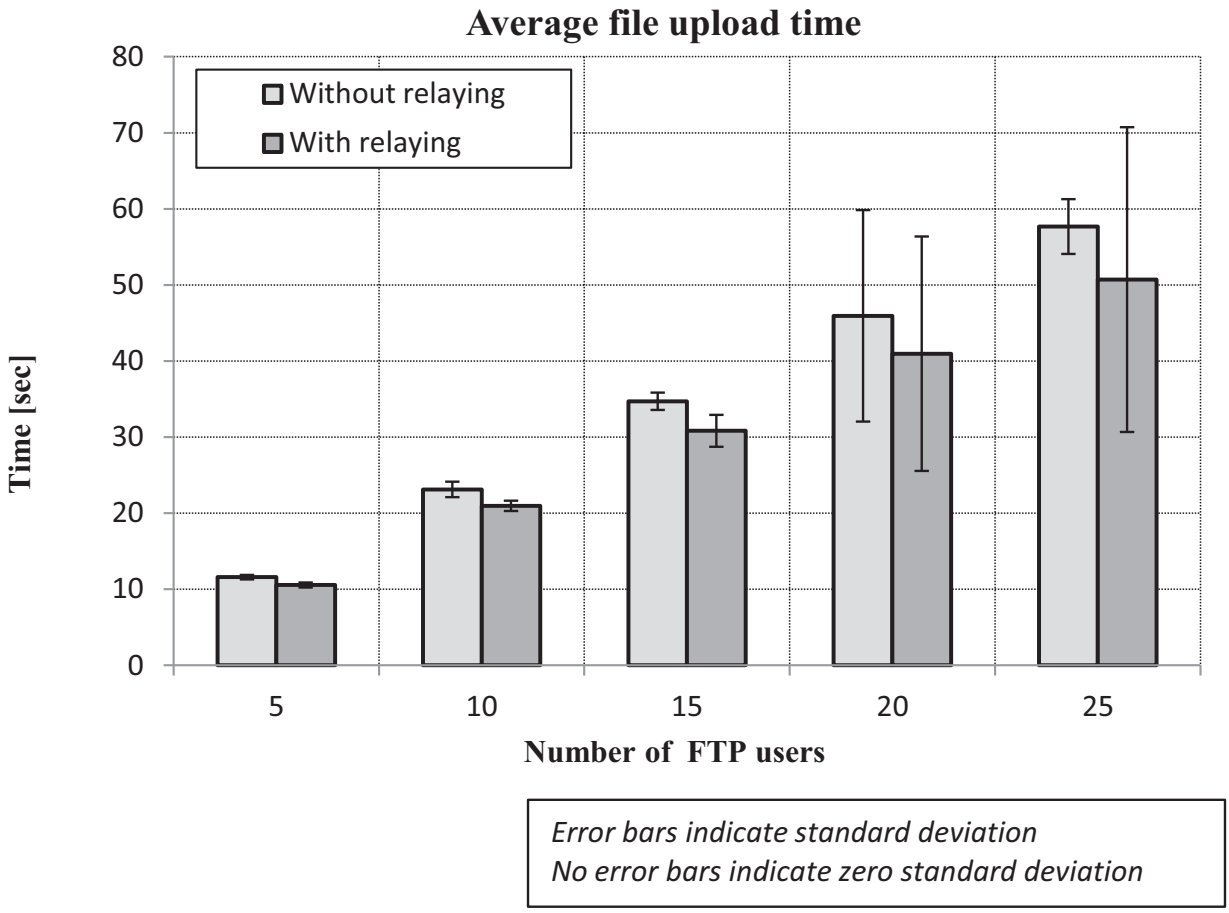

Figure 4.14: Average file upload time with and without RN for FTP traffic

4.6.2 M2M Traffic Aggregation and Multiplexing

The design and implementation of the M2M data traffic aggregation and multiplexing scheme into the OPNET simulation environment is evaluated with the help of two categories of scenarios. In the 1st category, the location of the RN is very close to the DeNB and corresponds to an MCS of 26. At this location, the TBS for RNs with $N_{\max}$ PRBs would be much larger than the TBS achievable at a location further away from the DeNB. The TBS with 5 PRBs would be 469 bytes as in Table 4.2. In the 2nd category of scenarios, the RN is moved slightly away from the DeNB and this location corresponds to an MCS of 16. As a result, the maximum TBS for the 5 PRBs of RN would be smaller (225 bytes) as compared to the TBS for the location in the 1st category. The multiplexer is designed to multiplex packets into large IP packets by considering the available TBS and overhead size. The maximum number of PRBs that can be allocated to an RN, $N_{\max}$ is fixed. The scenario specific simulation parameters are given in Table 4.4, whereas the cell loads for various number of M2M devices are given in Table 4.5.

Scenario parameter	Setting
RN location MCS	26
Maximum PRBs for RN	5 PRBs
M2M devices	1000, 2000, ..., 15000

Table 4.4: Scenario specific simulation parameters for RN near the DeNB

M2M devices	Load
1000	0.07
2000	0.14
3000	0.20
4000	0.27
5000	0.34
6000	0.41
7000	0.47
8000	0.54
9000	0.61
10000	0.68
11000	0.74
12000	0.81
130000	0.88
140000	0.95
150000	1.01

Table 4.5: Scenario specific cell loads with RN near the DeNB

The categories are subdivided into three subcategories of scenarios. In the 1st subcategory, the M2M traffic is served by the RN without multiplexing the data traffic. In the 2nd subcategory, the M2M traffic is multiplexed at the RN but without the timer expiry function. The packets are multiplexed only if the buffer size along with the overhead size reaches the achievable TBS of $n_{\max}$. In the 3rd subcategory, the timer starts at the arrival of packets into the RN buffer at the Un PDCP layer. If the maximum waiting time for multiplexing is exceeded, then the multiplexing of already arrived packets is performed even though the achievable TBS size has not been reached. All the subcategories of scenarios consist of subscenarios with different M2M traffic loads. In the 1st subscenario for all the subcategories, the number of M2M devices is 1000. In the successive subscenarios, the number of devices is increased by 1000. The error bars in graphs represent the standard deviation of the results.

Figure 4.15 depicts the average number of PRBs used by the RN for M2M de-

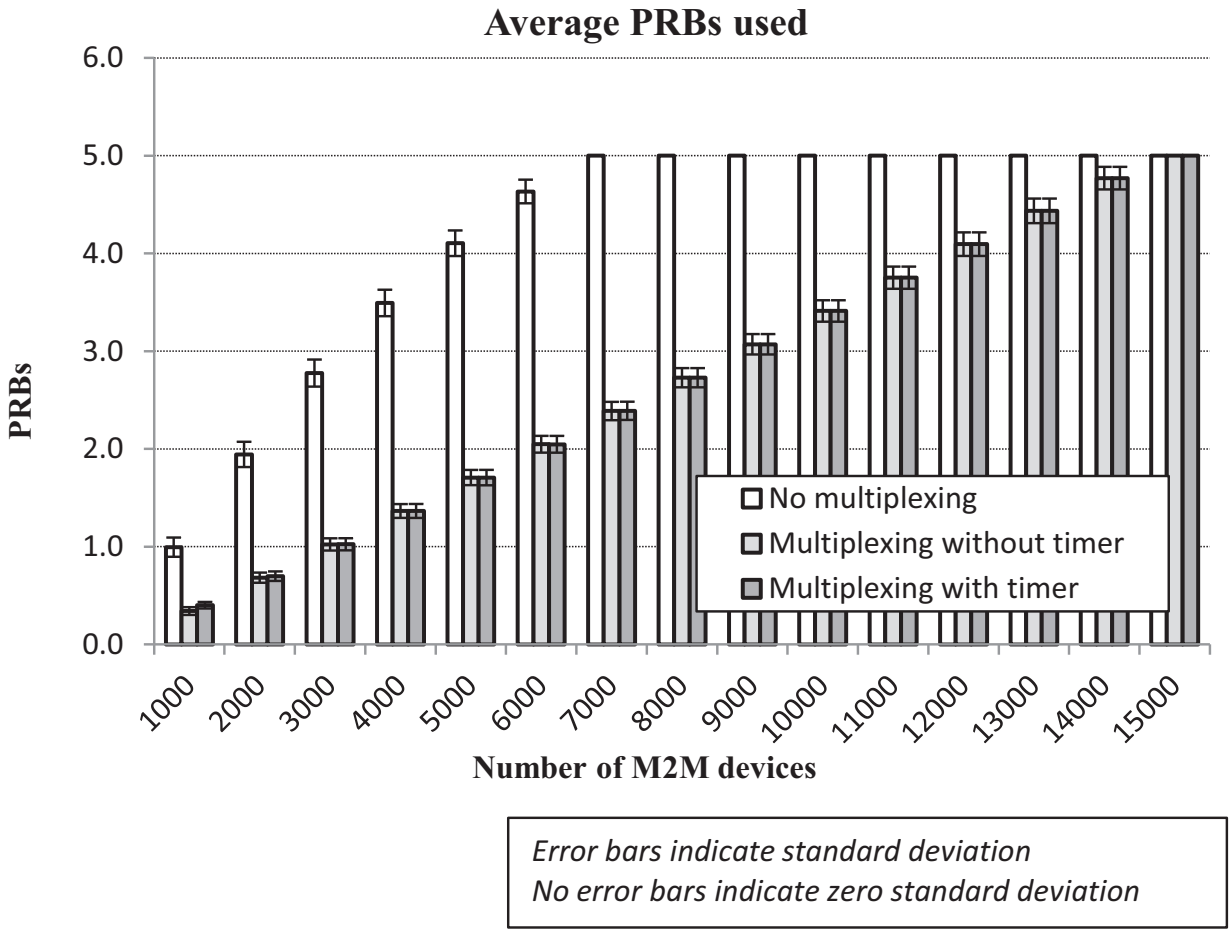

Figure 4.15: Average Uu PRBs usage with RN near the eNodeB

vices in various subscenarios. In all the three subcategories i.e. no multiplexing, multiplexing without timer and multiplexing with timer, the results are expected to show an increase in the number of PRBs used by the RN until the maximum number of PRBs, $N_{\max}$ is achieved. In low load scenarios where less than 7000 M2M devices are served by the RN, the average usage of PRBs increases with increasing the traffic load. The PRB usage is higher in case of RN with no multiplexing because the packets at the Un PDCP layer of the RN are going through the Uu air interface in such a way that each packet is accompanied by UDP overhead, IP overhead, PDCP overhead and RLC overhead. For broadband applications such as video surveillance or FTP upload, this amount of overhead can be considered as acceptable. However, for narrowband M2M traffic, this overhead would consume most of the available resources. Therefore, the aggregation and multiplexing scheme at RN results show that the radio resources could be significantly preserved. The results for multiplexing without timer are even better than the case of multiplexing with timer due to fact that without timer, the multiplexer waits for

sufficient arrivals. The goal of timer is to avoid long delays if the packets in the buffer are sufficient enough for filling the complete TBS. The timer is effective in in low load scenarios. In high traffic load, the timer is not required because sufficient data packets are always expected to be available in the buffer for multiplexing. In high load subscenarios, the average PRBs required by the RN reach $N_{\max}$. As a result, the packets have to wait for longer due to the unavailability of the resources. However, due to multiplexing of data traffic, the results for other categories show that even up to 1400 M2M devices can be served efficiently without the shortage of PRBs (Figure 4.15). When the number of devices is raised to 1500, the average PRB usage cannot increase and the multiplexing scheme fails to provide QoS effective service (as in Figure 4.16).

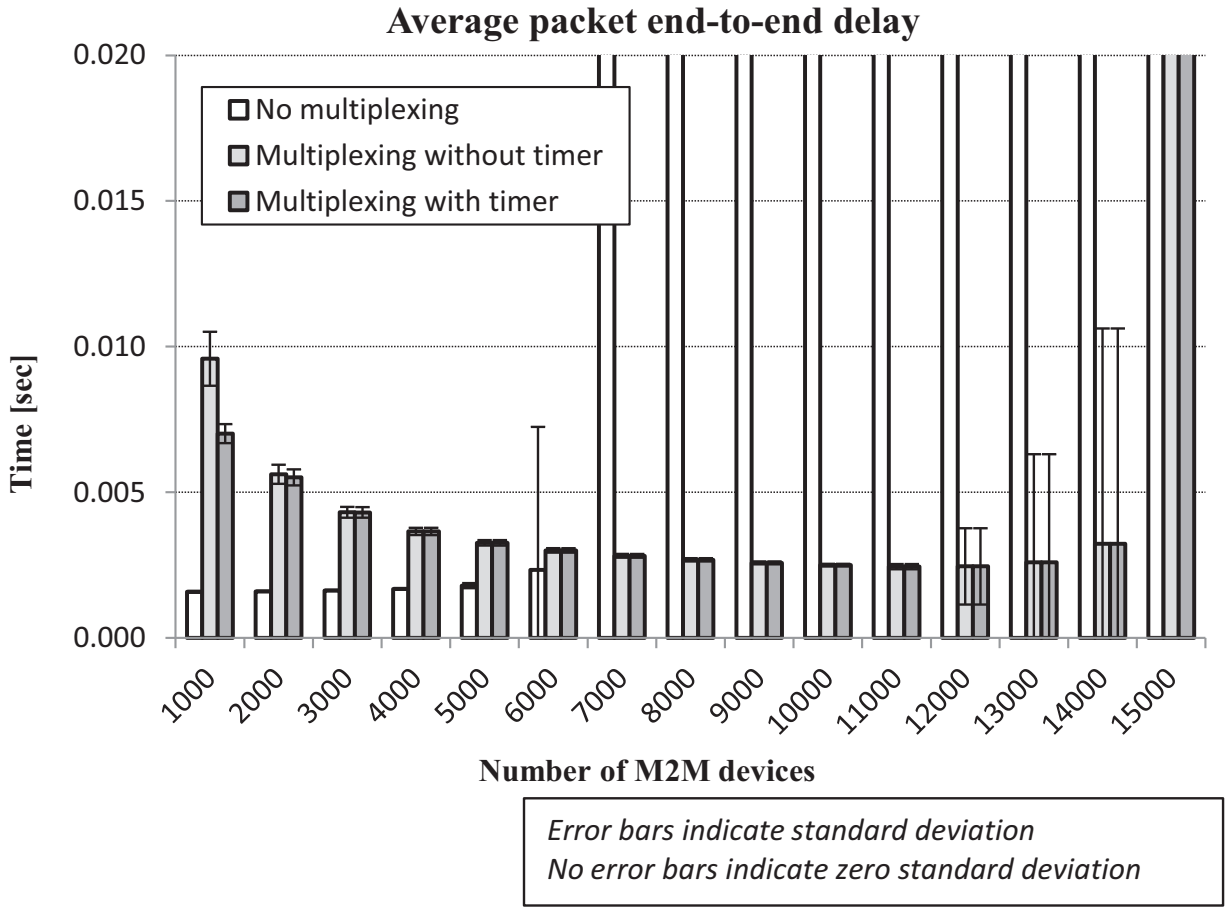

Figure 4.16: Average M2M packet end-to-end delay with RN near the eNodeB

As described, the QoS performance can be improved by the multiplexing scheme as shown in Figure 4.16. The performance of M2M traffic in terms of end-to-end delay can only be according to the QoS requirements if the traffic load is low. In low load subscenarios, the QoS performance with no multiplexing at the RN is

better than the performance of multiplexing case. The reason for this is the fact that the multiplexer waits for the packets until $n_{\max} - overhead^2$ bytes arrive at the buffer.

In case of the timer, the waiting time does not exceed $T_{\max}$. Therefore, the average end-to-end delay with timer is lower than without timer case. The timer plays a significant role only in low load scenarios. When the load is high, the RN is unable to fulfill the QoS requirements as the traffic end-to-end delay is very high (not in the scale of Figure 4.16). The multiplexer helps in improving the QoS performance significantly as clearly illustrated in Figure 4.16. The only subscenario, where the number of M2M devices is 15000, the multiplexing scheme at the RN also fails to achieve the end-to-end QoS requirements. With 15000 M2M devices, the load increases beyond 1 and makes the system unstable. -

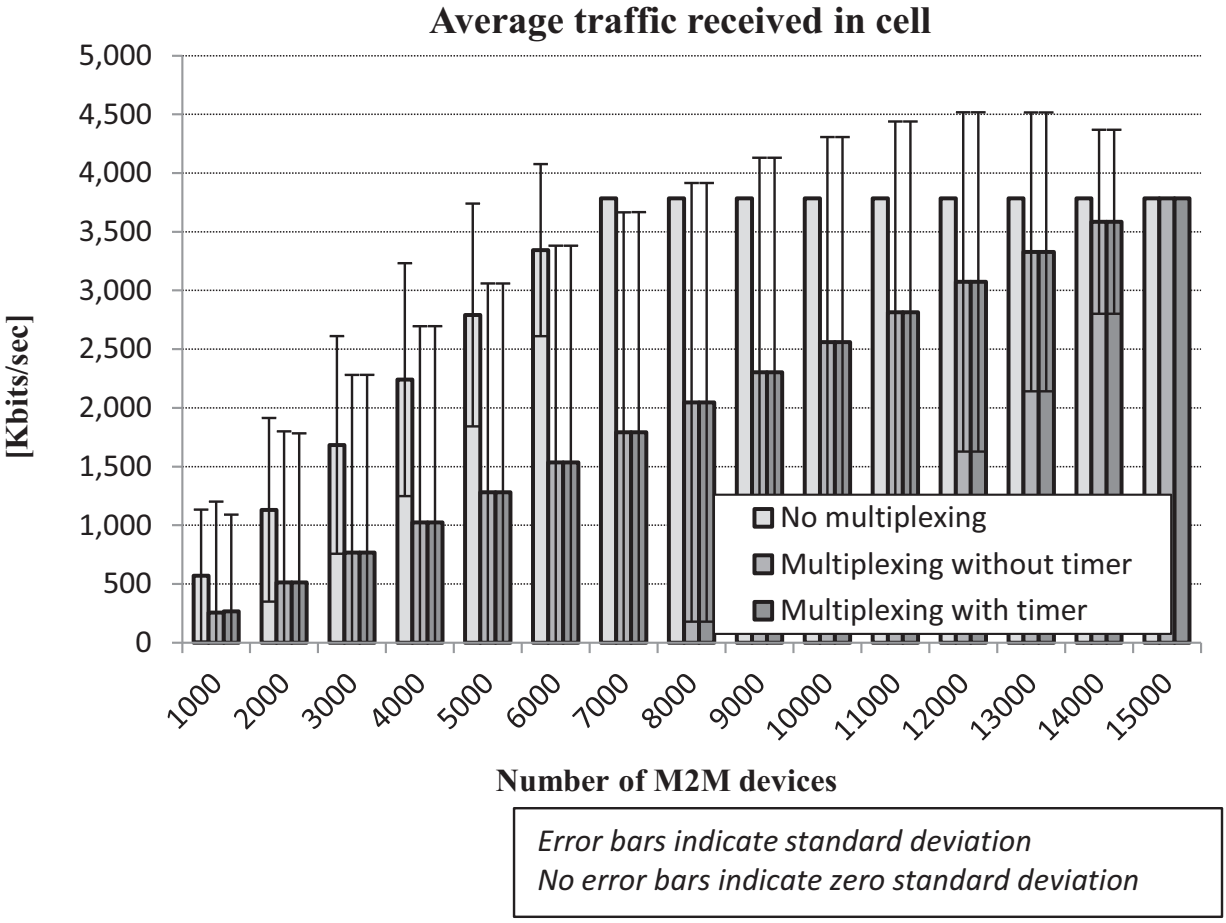

Figure 4.17: Average traffic received in cell with RN near the eNodeB

The received cell traffic is determined by the traffic sent towards the DeNB from the RN and illustrated in Figure 4.17. The low load cell traffic for the scheme

[2]overhead of GTP, UDP, IP, PDCP, RLC, MAC and PHY layers of Un interface of RN

without multiplexing is higher than the multiplexing schemes with and without timer. However, increasing the load to 7000 devices or more does not result in raising the received cell traffic. In case of multiplexing, the increase in load of up to 1400 results in the rise of traffic received in cell. If 15000 devices are deployed, then the RN load is higher than the cell capacity.

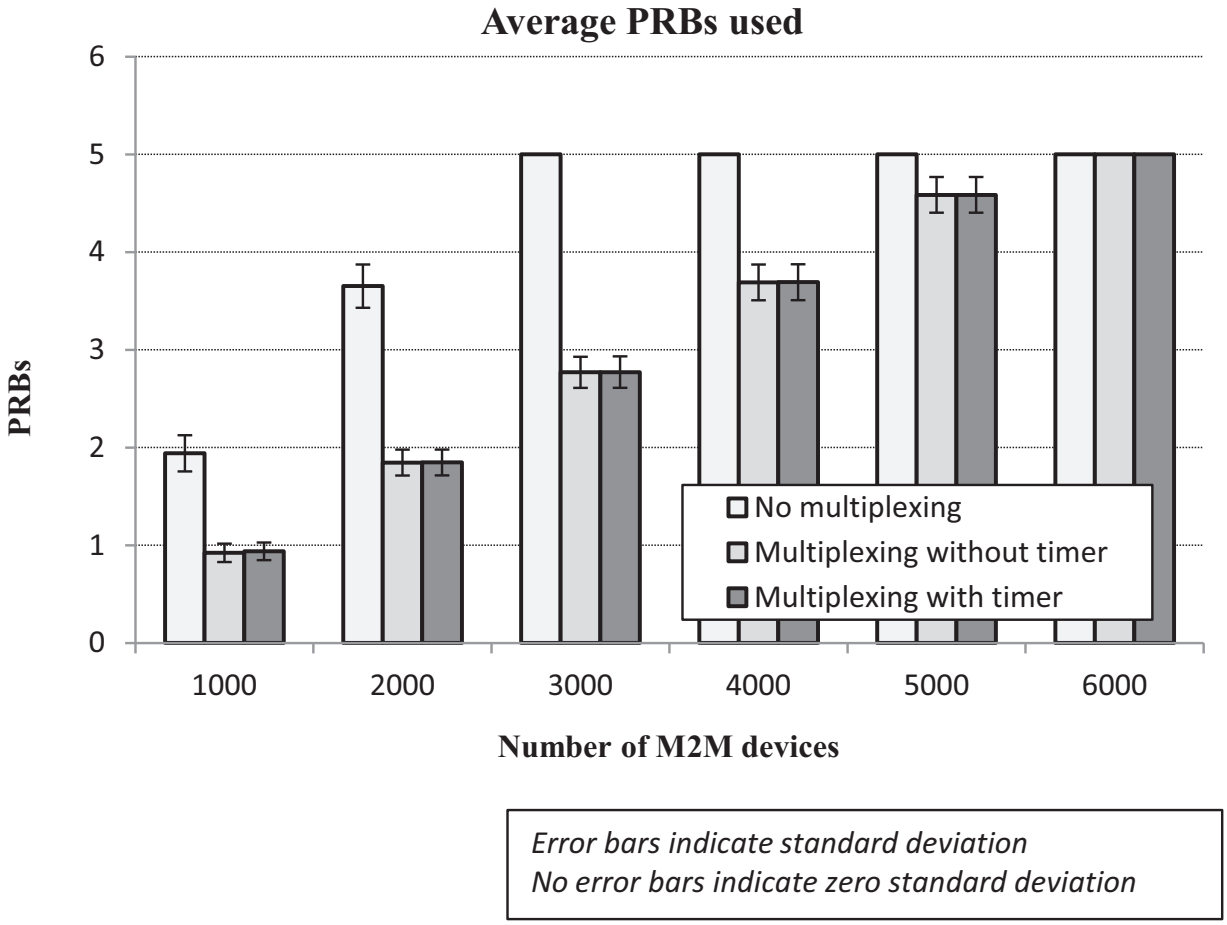

Figure 4.18: Average Uu PRBs usage with RN away from the eNodeB

Scenario parameter	Setting
RN location MCS	16
Maximum PRBs for RN	5 PRBs
M2M devices	1000, 2000, ..., 6000

Table 4.6: Scenario specific simulation parameters for RN away from the DeNB

In the 2nd category of simulation scenarios, the RN is moved away from the cell center. This location of the RN corresponds to MCS 16. The multiplexer is designed to adapt according to the channel conditions of the RN and the size of

M2M devices	Load
1000	0.18
2000	0.37
3000	0.55
4000	0.74
5000	0.92
6000	1.11

Table 4.7: Scenario specific cell loads with RN away from the DeNB

the multiplexed IP packet is in accordance with the TBS achievable. Since the capacity of a PRB is reduced when the channel conditions are poor, therefore, Figure 4.18 shows that it is not possible to serve even 3000 M2M devices without multiplexing at the RN. The whole resources are used up if there are 3000 M2M devices in the RN coverage area. With multiplexing, service could be provided to up to 5000 devices. The impact of the timer is only significant if there are 1000 devices. This implies that the timer is effective in low load cases. When 6000 devices are deployed in the RN coverage area, then even the multiplexing of traffic is unable to handle the traffic as the load has been increased beyond the capacity. The scenario specific simulation parameters are given in Table 4.6, whereas the cell loads for various number of M2M devices are given in Table 4.7.

The results for average packet end-to-end delay of the M2M traffic is depicted in Figure 4.19. The result for low load subscenarios, such as 1000 and 2000 devices show that multiplexing is not required for low loads. The end-to-end delay results in these subscenarios are better than the results for multiplexing with or without timer subscenarios. The role of the timer in reducing the multiplexing waiting time is also visible in the low load. However, when the number of devices are 3000 or more in the cell, the RN without multiplexing is unable to serve according to the QoS requirements of M2M traffic. The end-to-end delay results for these subscenarios are beyond the scale in Figure 4.19. When the number of devices reach 6000, even the multiplexing at the RN is unable to handle the M2M traffic anymore.

The results for received cell traffic are depicted in Figure 4.20.The traffic received from the RN with no multiplexing is higher than from the RN with multiplexing in the given scenarios, but this is due to the fact that all the data packets are accompanied by large overheads. The multiplexing of packets at the RN results in low data traffic requirements. At the point where the load exceeds the maximum possible cell capacity, the cell is not capable of fulfilling the QoS requirements to devices.

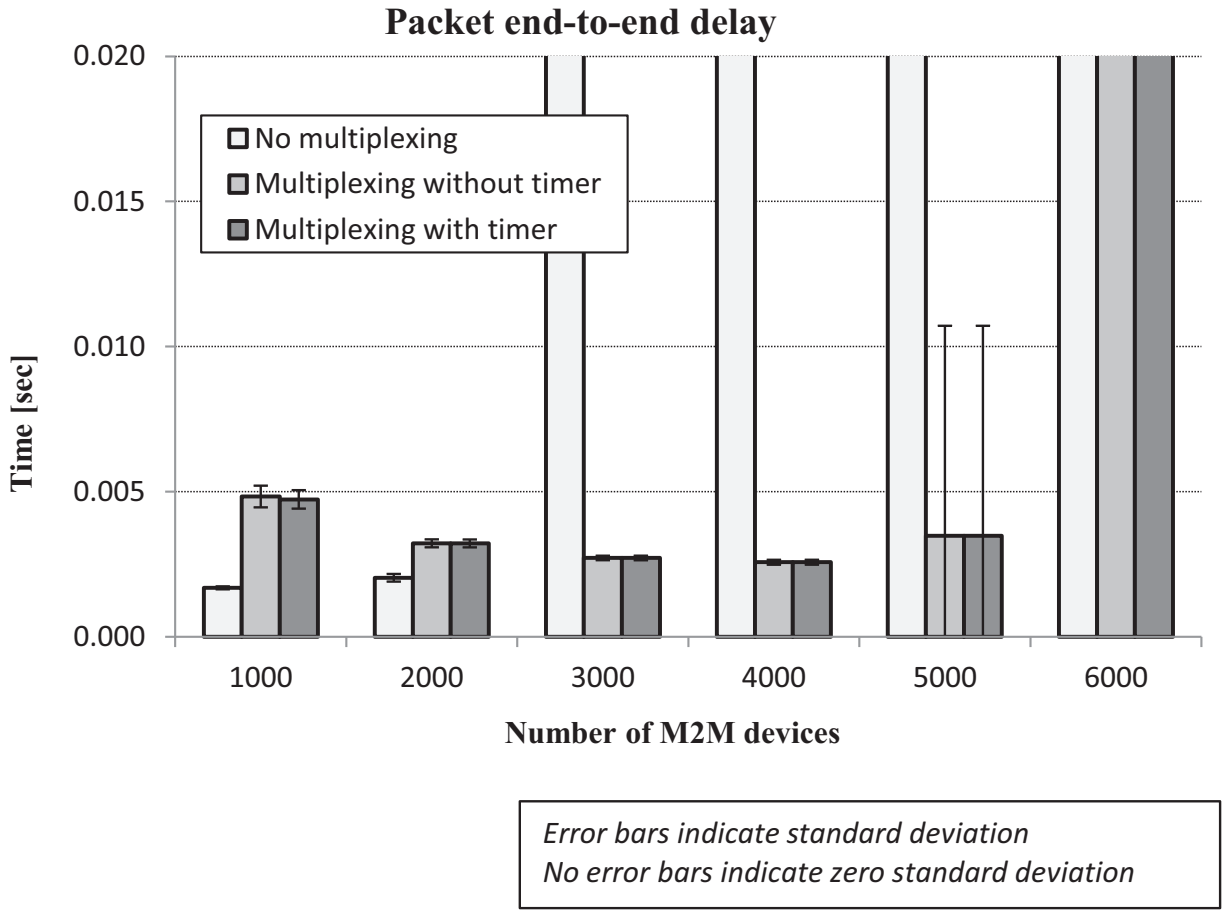

Figure 4.19: Average M2M packet end-to-end delay with RN away from the eNodeB

4.6.3 Impact of M2M Relaying on Regular Traffic

The impact of M2M communication on regular LTE-A data traffic can be adverse as depicted in section 3.6 (for impact on LTE traffic). The RN based data aggregation and multiplexing framework can be utilized for minimizing the influence of M2M traffic on regular LTE-A traffic. This beneficial effect is illustrated in this section with help of simulation scenarios where regular LTE traffic is also deployed in the eNodeB coverage area along with the M2M traffic. The regular LTE traffic consists of 15 FTP users. The M2M traffic is served by the RN with multiplexing of data and the waiting time of packets is managed by the timer. The location of the RN corresponds to an MCS of 22. The scenarios are divided into two categories. In one category, the M2M traffic is served directly by the eNodeB along with the FTP users. In the 2nd category, the M2M traffic is served by the RN and the FTP users are obtaining radio resources from the eNodeB directly. Both categories are further subdivided in subscenarios in such a way that the number of M2M devices is varied in each subscenario. In the 1st subscenario of both the

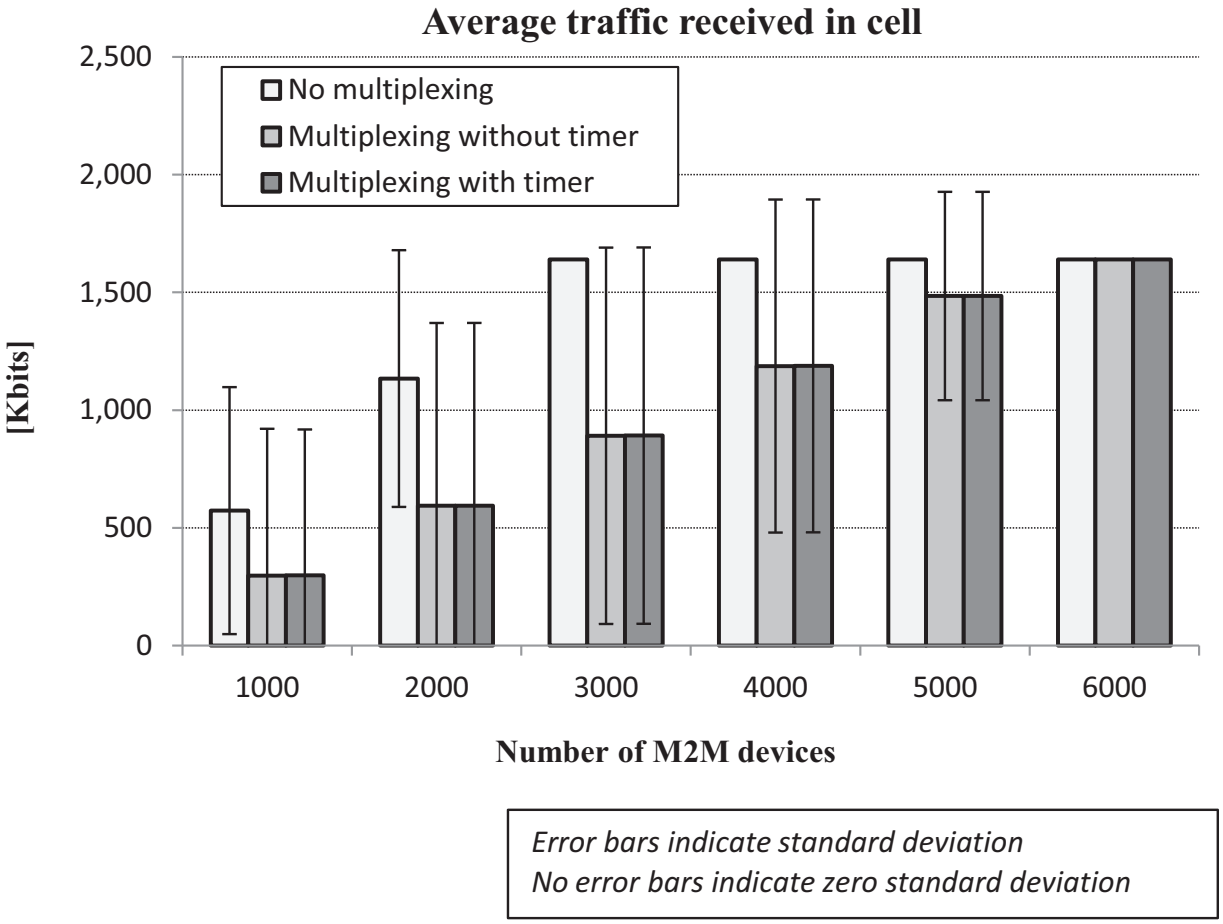

Figure 4.20: Average traffic received in cell with RN away from the eNodeB

categories, 1000 M2M devices are deployed. This number is increased by 1000 as compared to the previous subscenarios. The error bars in graphs illustrate the standard deviation of the results of subscenarios. The scenario specific simulation settings for determining the impact of M2M relaying on regular LTE-A traffic is given in Table 4.8.

Scenario parameter	Setting
RN location MCS	22
Maximum PRBs for RN	5 PRBs
M2M devices	1000, 2000, ..., 5000
FTP UEs	15

Table 4.8: Scenario specific simulation parameters

The average number of PRBs used by all the devices and UEs in all the scenarios is equal to 25 because the best effort traffic is present in the cell (Figure 4.21). Any

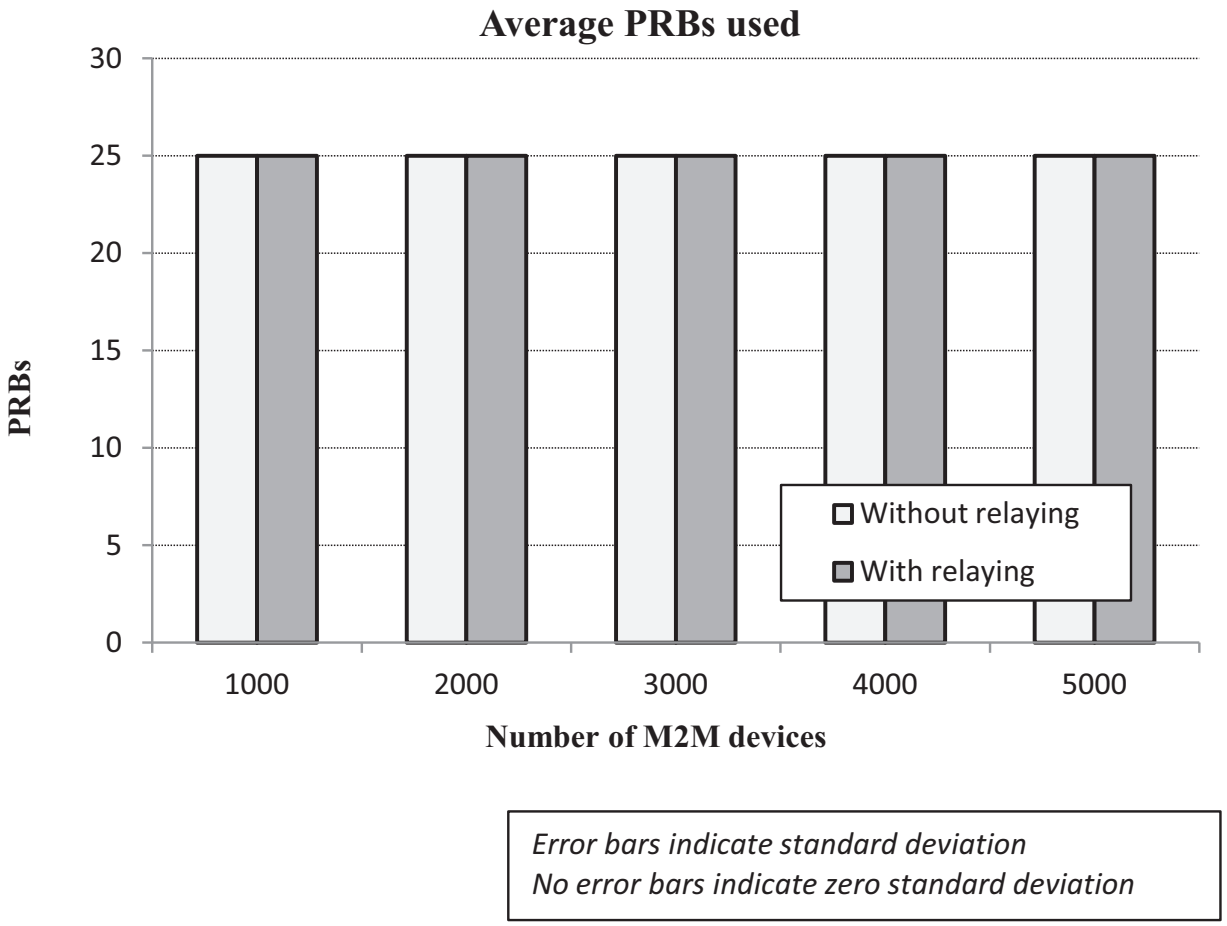

Figure 4.21: Average PRB usage with and without RN

unused resources are utilized by the FTP users. Therefore, the PRBs are utilized completely in all the subscenarios.

The results for average packet end-to-end delay of M2M traffic are shown in Figure 4.22. In low load subscenarios, the eNodeB is capable of directly serving the M2M traffic efficiently, without having to deal with the relaying delay. Therefore, the average end-to-end delays for the category of M2M service through the RN are higher than without RN. However, when the number of devices reaches 5000, the eNodeB is unable to serve the M2M traffic with the fulfillment of QoS requirements anymore. However, the RN based M2M service can still be accomplished in an effective manner.

The average FTP file upload time for all the subscenarios is depicted in Figure 4.23. Increasing the load results in higher file upload times in all the subscenarios. But one can clearly notice that the deployment of the RN for M2M data traffic is proving to be greatly beneficial for the FTP traffic in the cell. The impact of M2M traffic on LTE-A regular traffic has been minimized under the given

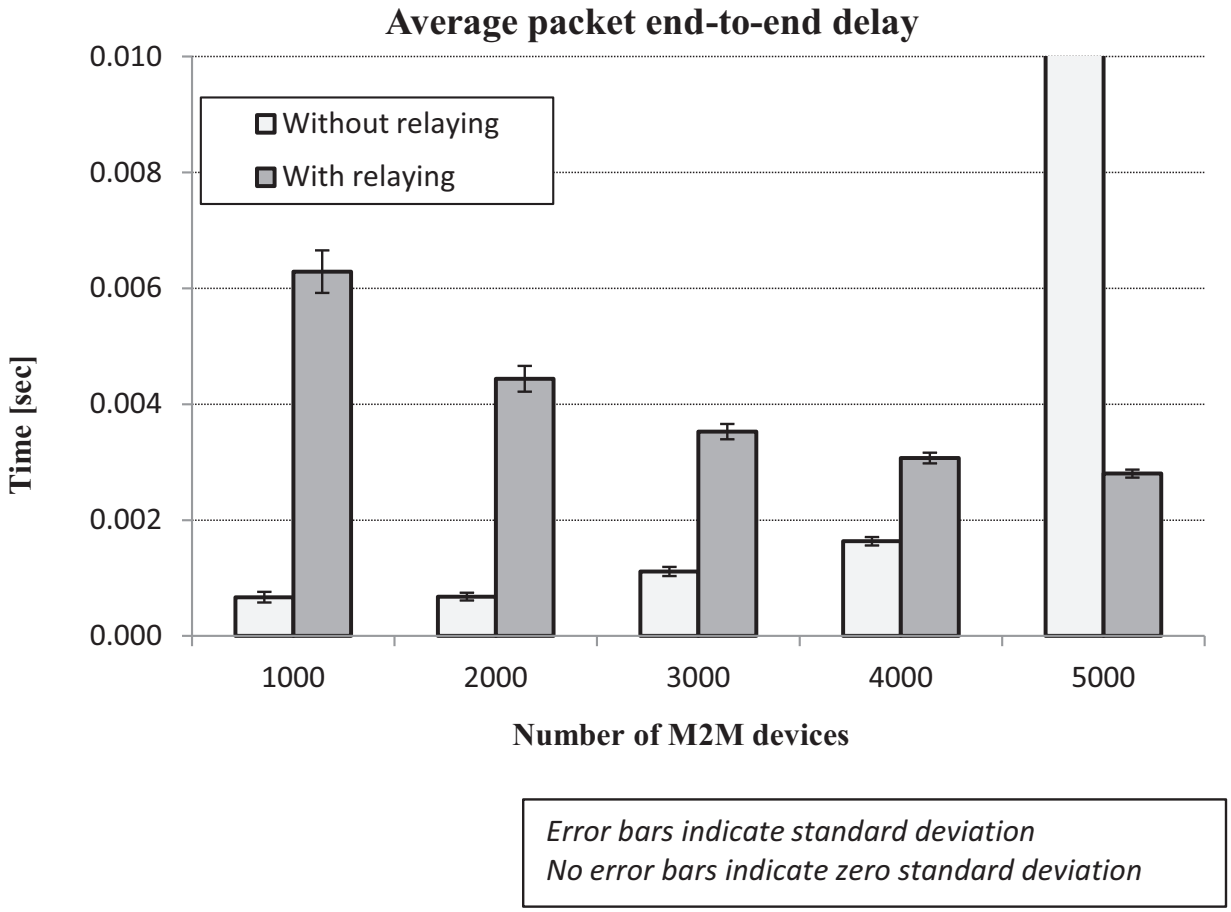

Figure 4.22: Average M2M packet end-to-end delay with and without RN

simulation setup significantly.

The average cell traffic performance of the subscenarios in both the categories is illustrated in Figure 4.24. The average traffic received in cell increases with increasing the number of M2M devices for the results in the category of M2M traffic over RN. However, the cell traffic decreases if the load is increased in the results under the category of direct M2M service by the eNodeB. The reason behind this performance degradation is that the FTP traffic can be considered as broadband traffic. The LTE-A air interface resources are efficiently utilized if the broadband traffic is transmitted. However, increasing the narrowband M2M traffic in the cell results in reduced cell traffic because the PRBs are allocated to small size data messages from individual devices. Whereas in case of RN multiplexing, the PRBs are shared by data emerging from different devices.

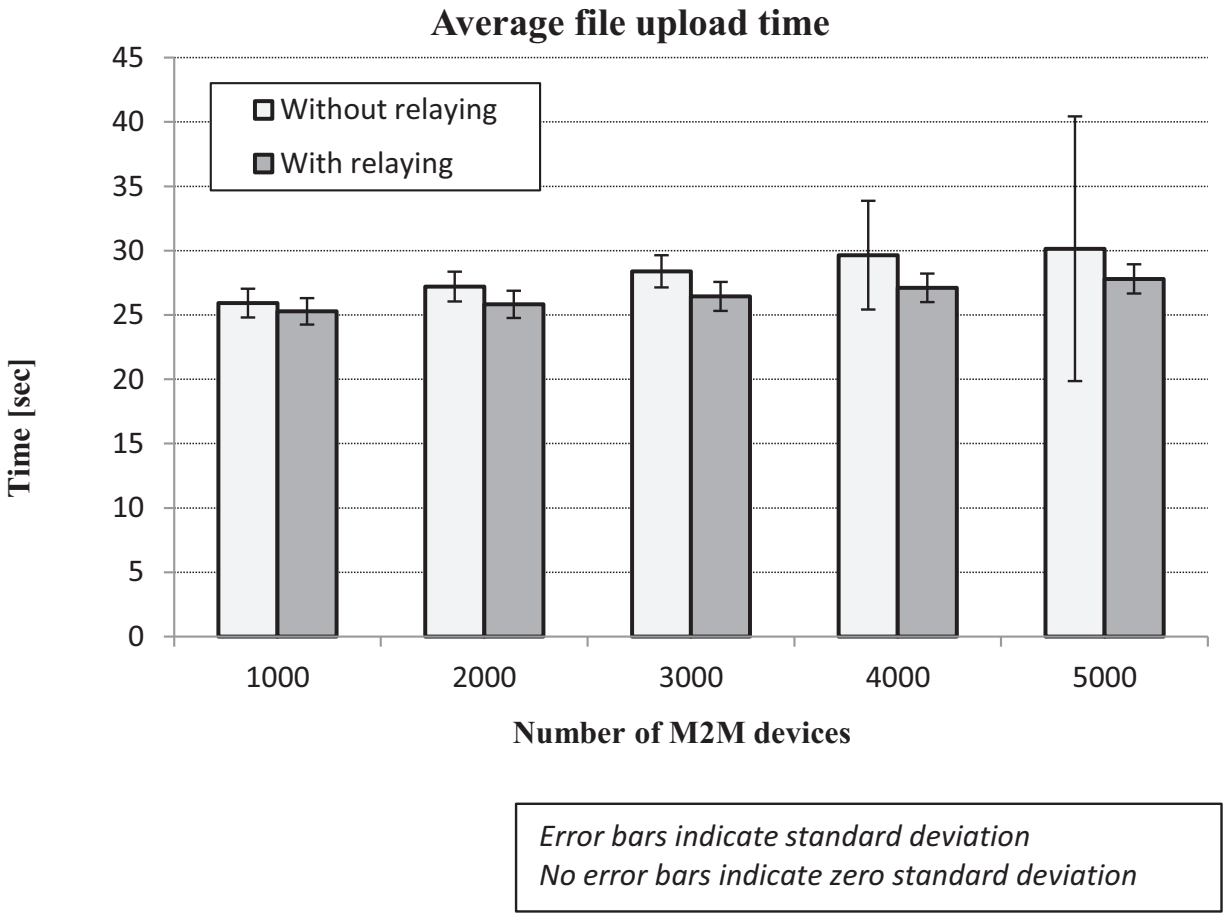

Figure 4.23: Average FTP file upload time with and without RN

4.7 Conclusion

This chapter provided a detailed overview of relaying system functionalities. The RN was utilized as a low cost solution for the service of M2M traffic without introducing complexity issues into the communication system. An RN based framework was presented for the aggregation and multiplexing of data before transmission over the Un interface.

The RN based data traffic and QoS performance was evaluated for the cell edge users. The performance of cell edge users significantly improved with the RN. The end-to-end delay times were reduced due to the availability of better channel conditions with the deployment of RNs. Dead spots such as basements can also be served in this manner.

The aggregation and multiplexing scheme was evaluated with two different locations for the RN in the cell for M2M data traffic. In both the cases, the performance of M2M traffic in terms of end to end delay and cell traffic improved in high load

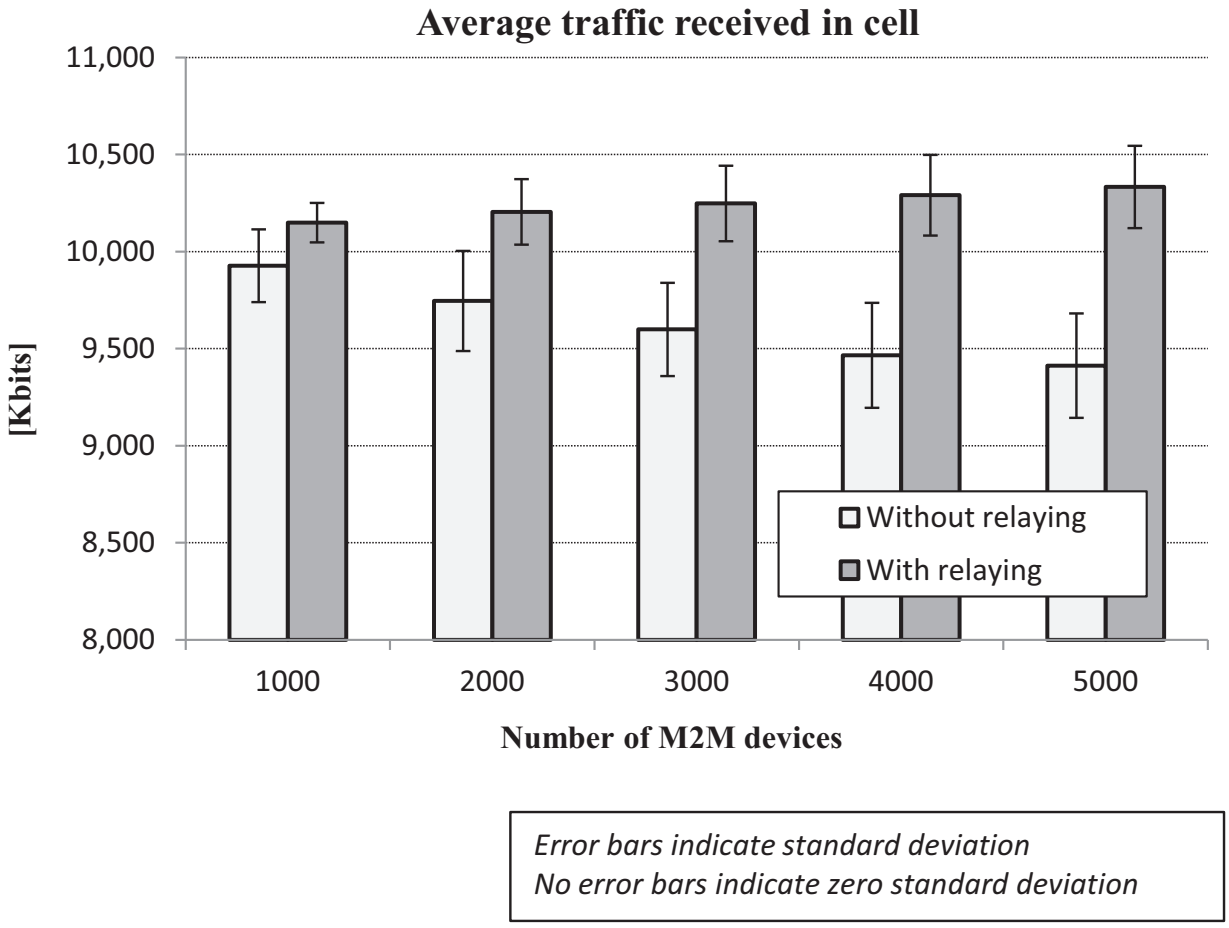

Figure 4.24: Average traffic received in cell with and without RN

subscenarios. The future M2M traffic is expected to be based on a huge number of devices. The proposed scheme can be seen as a viable solution for capacity problems faced by the network operators.

The impact of the proposed RN based multiplexing scheme on the QoS provisioning to regular LTE-A traffic was also investigated. The FTP UEs were requesting resources directly from the eNodeB and the M2M devices were served by the RN. The results showed that the proposed framework can significantly minimize the impact of M2M traffic on regular LTE-A traffic.

5 Results Comparison for Relay Node

An analytical model of a system is developed with the primary goal of evaluating the system performance [Zak12]. Analytical modeling is required for the evaluation and understanding of complex systems to avoid expensive field tests and exhaustive simulations. Analytical modeling helps in saving time and resources. In many cases, it is not even possible to perform field experimentation. For instance, the performance evaluation of a metropolitan area network covering a city would require significant time, power and other resources. Simulation models are also usually very complicated with huge computational times for the evaluation of system performance. Understanding the complexities of a simulated system requires additional time. The motivation behind analytical modeling is the approximation of the behavior of a system mathematically and reducing the time and effort required for processing.

In this thesis, an analytical model for the performance evaluation of the Relay Node (RN) data traffic aggregation and multiplexing scheme is presented. The model is based on the work of [JSC11] for multiplexing of small Common Part Sub-layer (CPS) packets into large ATM cells with the AAL2 layer multiplexer. An ATM cell is made up of a 5 byte header and a 48 byte payload. Among the 48 bytes of payload, 1 byte is reserved for the start field. The rest of the 47 bytes are available for multiplexing the short CPS packets. The model features two queues, the multiplexing queue and the transmission queue. At the multiplexing queue, the short CPS packets wait for multiplexing until either the number of packets in the queue is sufficient to fill the ATM payload or a timer expires. In case, the timer comes into play, then the rest of the unused bits in the ATM payload are zero padded. The transmission queue is for the ATM cells with a maximum of r CPS packets in each ATM cell. The maximum number of packets r in an ATM cell can be determined as

$$r = \left\lceil \frac{47}{l} \right\rceil \tag{5.1}$$

where l is the fixed size of a CPS packet.

The comparison of results from simulation and analytical models provides validation to the simulation results if the results are in agreement with each other. It is common to have slight differences in the results primarily due to the reason that the analytical model is abstract and covers only the main features of a system. The simulation model is complex and it is very difficult to capture all the aspects of simulation in analytical model.

In order to further consolidate the notion that the simulation results are valid and acceptable, an abstract simulation model is also developed in this thesis, referred to as 'simple' simulation model. The purpose of the simple simulation model is to minimize the influence of complex simulation aspects and highlight the performance of the proposed scheme purely.

5.1 The Analytical Model

The approach in [JSC11] is analogous to the multiplexing scheme proposed in this thesis. The M2M data packets arrive at the Uu PDCP layer of the RN from various M2M devices. The arrival of a packet to the multiplexer at the RN PDCP layer sets a timer in such a way that the packet has to wait for a certain maximum duration of time $T_{\max}$ for the arrival of more packets to achieve maximum multiplexing gain. If the waiting time exceeds the timer expiry limit $T_{\max}$, the multiplexing process is initialized and the multiplexed data is sent to the Un PDCP buffer of the RN. If all the $N_{\max}$ PRBs are not required for transmission, the unused PRBs are allocated to other users requesting resources from the DeNB. In case, the maximum waiting time $T_{\max}$ is not exceeded, but the accumulative size of the packets in the buffer exceeds the TBS capacity ($n_{\max} - overhead$), then the multiplexing starts even if the timer has not expired. The overhead from GTP, UDP, IP, PDCP, RLC, MAC and PHY layers of the Un interface of the RN have to be considered while determining the TBS capacity [Sem14]. The size of $n_{\max}$ depends on the TBS which depends on the channel conditions of the backhaul air interface. The maximum number of packets r that can be multiplexed into a large IP packet can be expressed as

$$r = \left\lceil \frac{n_{\max} - overhead}{l} \right\rceil \tag{5.2}$$

where l is the fixed size of the packet arriving at the RN Uu PDCP layer.

If there are k M2M packets in the PDCP buffer, then until the waiting time reaches $T_{\max}$ or the buffer size reaches $n_{\max} - overhead$, the value of k is such that $1 \leq k < r$. The arrival rate of the M2M data packets is considered to be exponentially distributed with a rate of λ. The total waiting time of k^{th} packet arrival time

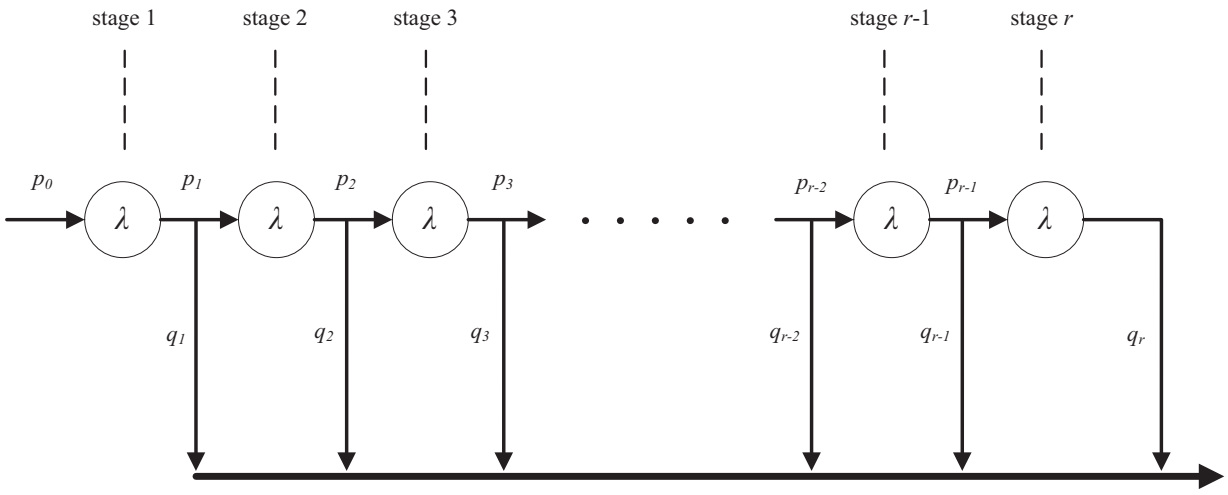

Figure 5.1: r-stage Coxian distribution

is Erlangian [JSC11]. The probability of the next arrival before multiplexing with k packets already in the buffer is p_k, whereas the probability that the multiplexing process starts after the arrival of k^{th} packet is q_k such that $p_k = 1 - q_k$. This process can be modeled as an r-stage Coxian process and is depicted in Figure 5.1. The probability of the next arrival with k packets in the buffer can be determined using (5.3)

$$p_k = \begin{cases} 1, & k = 0 \\ 1 - \left(\sum_{j=0}^{k-1} \frac{e^{-\lambda T_{\max}} (\lambda T_{\max})^j}{j!} \right), & 1 \leq k < r \\ 0, & k \geq r \end{cases} \tag{5.3}$$

The probability of an arrival with 0 packets in the buffer, p_0, is always 1 because it is not possible that the multiplexing would start without any packet in the buffer. To multiplex packets, at least one packet in the buffer is required. Therefore, q_0 is always 0. After the arrival of 1^{st} packet, the system enters stage 1 of the r-stage coxian process shown in Figure 5.1 (i.e. there is a packet in the buffer). In this stage, it is possible that a 2^{nd} packet arrives in the buffer before the timer expires with a probability p_1. But it is also possible that there are no arrivals after the 1^{st} one and the timer expires to start the multiplexing process, the probability of which is q_1. Now if a 2^{nd} packet arrives before the timer expiry, the process enters stage 2. In stage 2, there are two possibilities, either another packet arrives (p_2) or the timer expires after 2^{nd} arrival (q_2). This can go on until r arrivals, which is the maximum number of packets possible in the buffer. At stage r, the only possibility

is that the multiplexing should start without waiting for any further arrivals. Thus, $p_r = 0$ and $q_r = 1$.

The Laplace-Stieltjes Transform (LST) of the Probability Distribution Function (PDF) of the r-stage coxian process can be expressed as

$$A(s) = \sum_{k=1}^{r} \phi_k \left(\frac{\lambda}{s+\lambda} \right)^k \tag{5.4}$$

where ϕ_k is the path probability and can be expressed as

$$\phi_k = q_k \prod_{j=0}^{k-1} p_j \tag{5.5}$$

where $p_0 = 1$. Equation (5.5) implies that $\phi_k = p_1 p_2 ... p_{k-1} q_k$.

5.2 The Simulation Model

Multiplexing transition probabilities in the simulation model are determined using a packet counting method. p_k is the probability of $(k+1)^{\text{th}}$ arrival before the multiplexing process with k packets in the buffer. In order to determine values of p_k in the simulation model, the maximum number of packets that can be multiplexed is considered as r. So k characterizes the number of arrivals in the buffer and can have any value from the range 1 to r such that $1 \leq k < r$.

To determine the a-posteriori probabilities of p_k for different values of k, a record of the number of times, an aggregated packet contained a certain number of packets is maintained. For this purpose, a counter for each stage of the aggregation buffer k (or in other words, a count of the number of times, the r-stage coxian process enters a particular stage k is recorded). These counters are defined as C_k for all $1 \leq k < r$, such that C_k counts the number of times the aggregation packet consisted of 1 or more packets (i.e. the number of times, the r-stage coxian process entered stage k).

Probabilities p_k can be determined on the basis of values of specified counters C_k. For instance, in order to find the value of p_1, the value of C_2 has to be taken into account. This gives the probability of 2_{nd} arrival with respect to C_1 (the total number of arrivals). So, p_k can be calculated as in (5.6).

$$p_k = \begin{cases} 1, & k = 0 \\ \frac{C_{k+1}}{C_1}, & 1 \leq k < r \\ 0, & k \geq r \end{cases} \tag{5.6}$$

The simulation parameters are given in Table 5.1.

Parameter	**Setting**
Cell layout	7 eNodeBs with single cell
System bandwidth	25 PRBs ($\approx$ 5 MHz)
Inter eNodeB distance	500 m
Max uplink transmit power	23 dBm
Channel models	Vienna simulator [IWR10]
Noise per PRB	-120.447 dBm
Noise floor	9 dB
Power Control	FPC, $\alpha = 0.6, P_0 = -58$ dBm
Timer expiry $T_{\max}$	9 msec
M2M traffic model	
Message size	Constant(29) bytes (including 28 bytes overhead of PDCP, RLC, MAC, PHY)
Message inter-transmission time	1 sec

Table 5.1: Simulation parameters

5.3 The Simple Simulation Model

Alongside the analytical model, a simple simulation model is developed to verify the simulation results. This is done to nullify the impact of the complex features of simulation model such as protocol stack, tunneling, wireless connectivity etc. Thus, the results of the simple simulation model would only illustrate the performance of the proposed scheme.

Figure 5.2: Simple simulation node model

The simple model is also developed in OPNET environment. The project model consists of a single node which includes four process models as in Figure 5.2. The process model labeled 'src' is the data traffic source. Instead of deploying users in simulation scenarios, the same kind of traffic is generated from 'src'. The process model 'relay' emulates the multiplexing functionality of the RN. If the multiplexing is turned on, then the packets would not be sent forward to the next process model 'denb' until either i) the buffer size reaches the TBS capacity for any given

MCS and number of PRBs or ii) the timer set for a specific duration expires. If the multiplexing is turned off, arriving packets are immediately sent to ‘denb’. In ‘denb’ process model, if multiplexing is turned on, the arriving packets are demultiplexed before sending them to the ‘sink’ process model. In case, multiplexing is turned off, the arriving packets are sent to ‘sink’ as soon as they arrive. The process model ‘sink’ is the module that destroys the arriving packets.

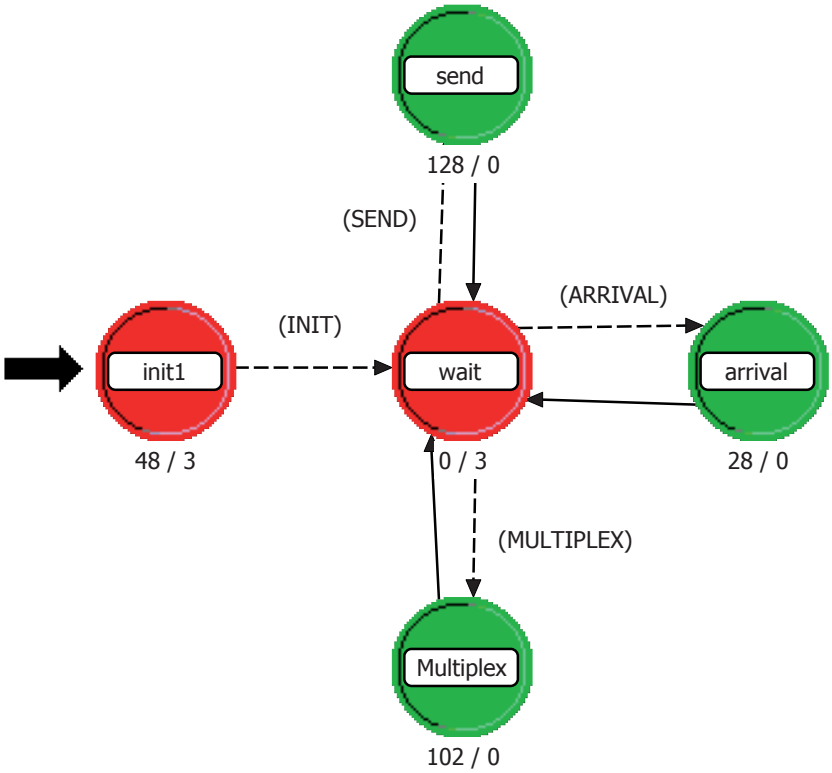

Figure 5.3: Simple simulation relay process model

The process model ‘relay’ is made up of states and transitions. The red colored states are unforced states (‘init1’ and ‘wait’); the transition from unforced state occurs only if a certain event occurs. The green colored states are forced states(‘send’, ‘arrival’ and ‘Multiplex’); the transition from forced state occurs as soon as the code in such states is executed. The starting phase of the process model is the ‘init1’ state (Figure 5.3)). In this state, the state variables are initialized once in a simulation run. The ‘wait’ state is where the process should stay while waiting for events to occur. In case of ‘ARRIVAL’ event, which means that a packet has arrived from ‘src’, the process goes to ‘arrival’ state. In case, the multiplexing is on, the arriving packet is stored in a packet list. If multiplexing if off, the packet is placed in a segmentation buffer. The event ‘MULTIPLEX’ occurs after every 1 ms. If there are any packets in the multiplexing buffer, then the timer is initi-

ated or decremented by 1 ms. If the buffer size has reach the TBS capacity, or the timer reaches 0, the data in the buffer is multiplexed into a large packet. The event 'SEND' occurs after every 1 ms as well. Any packets in the packet list (in case the multiplexing is off) or in the segmentation buffer (in case the multiplexing is on) are sent to the 'denb' process model in the 'send' state. Additionally, statistical results of multiplexing transition probabilities are also collected in the 'send' state. In the process model 'denb' (Figure 5.4), the reverse of the above mentioned course of action is performed and the packets are sent to the 'sink' process model.

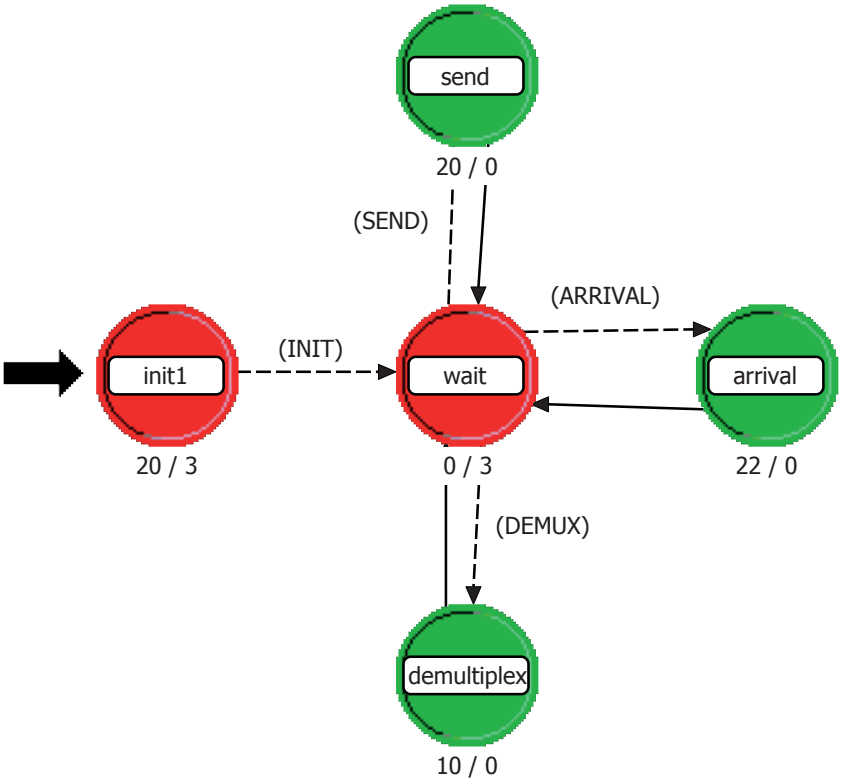

Figure 5.4: Simple simulation DeNB process model

5.4 Performance Evaluation

The performance evaluation of the data aggregation and multiplexing scheme for M2M traffic can be done by analyzing various parameters. Parameters considered in this thesis are *multiplexing transition probabilities*, *path probabilities*, *PRB utilization* and *multiplexing gain*. The RN is considered by the DeNB MAC scheduler as a GBR user, therefore the availability of radio resources ($N_{\max}$ PRBs) is ensured by the DeNB MAC scheduler for the RN. So multiplexed data can be considered to obtain radio resources immediately after the multiplexing process is performed.

Therefore, multiplexing is the primary factor that has an impact on the end-to-end delay performance.

A comparison of simulation results, analytical results, and the results obtained from the simple simulation model are given in Figure 5.5 to Figure 5.9 for multiplexing transition probabilities. The results for path probabilities are depicted in Figure 5.10 to Figure 5.14. These results are obtained for the case when the RN is placed at a location in the cell where an MCS of 16 can be achieved. The maximum number of RN PRBs, $n_{\max}$ is set to 5. This results in a TBS of 201 bytes (1608 bits). If the 44 bytes (352 bits) of overhead from lower layers of the RN Un is subtracted, the remaining 157 bytes (1256 bits) can be used for multiplexing the data packets. Since the size of each packet is kept as 29 bytes (232 bits) (as in Table 5.1), therefore, the maximum number of packets in a TBS can be 6 according to (5.1). The performance evaluation is achieved for low load scenarios of 100, 200, 300, ..., 900 and 1000 M2M devices served by the RN because the multiplexing delay plays no role in high load scenarios. The 95% confidence intervals for simulation results are provided in Appendix.

5.4.1 Multiplexing Transition Probabilities and Path Probabilities

The values of p_k for different values of k are determined for some of the low load scenarios using the simulation, analytical and simple models.

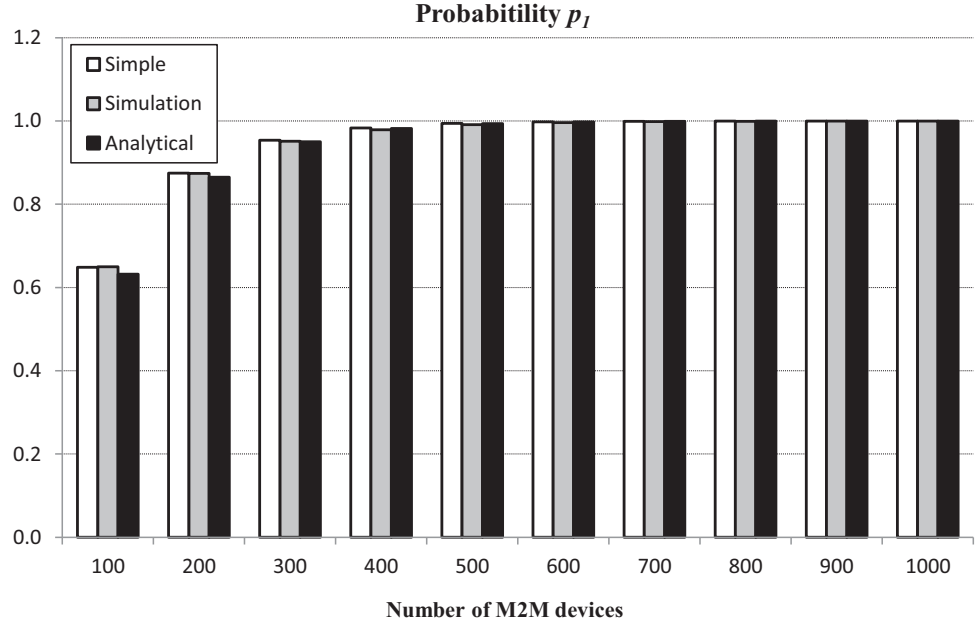

Figure 5.5: Comparison of probability p_1

p_1 is the probability of the arrival of 2nd packet before the multiplexing process starts. In case of 100 M2M users within the coverage area of the RN, the probability p_1 is around 0.6 for all the three models, i.e. simulation, analytical and simple model results as in Figure 5.5. This implies that the probability of timer getting expired before the 2nd arrival is approximately 0.4. Increasing the traffic load increases the probability p_1 in subsequent scenarios. In case of 700 or more M2M devices in the RN coverage area, p_1 is almost equal to 1.0 which indicates that at low load, the probability of multiplexing before 2nd arrival is very low.

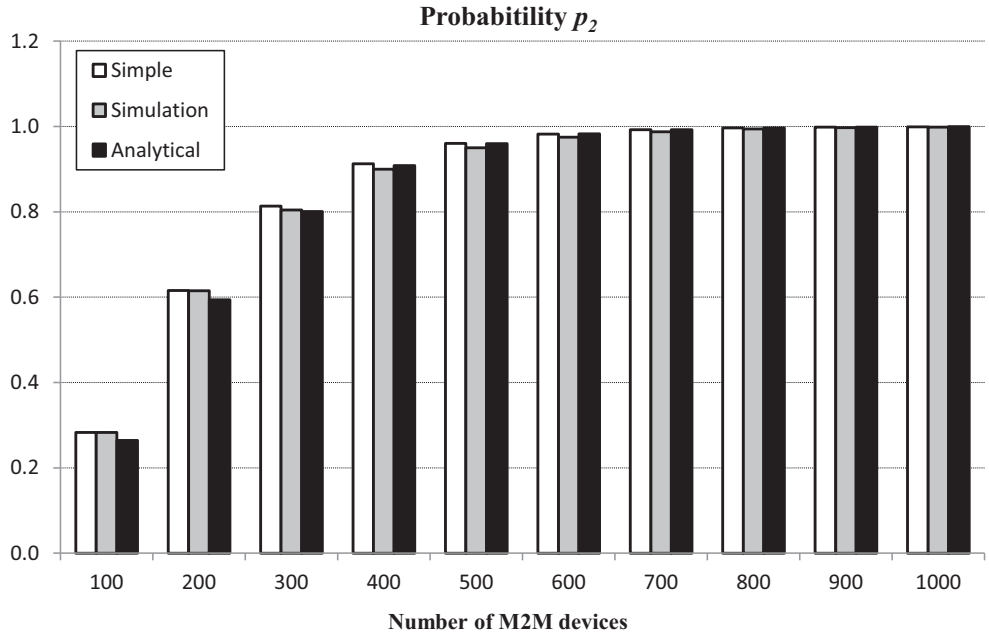

Figure 5.6: Comparison of probability p_2

Similarly, p_2 is the probability of the arrival of 3rd packet before the multiplexing process starts. The results for p_2 show that the arrival of 3rd packet is highly probable in high load scenarios as compared to low load scenarios (Figure 5.6). As expected, p_2 values are slightly less than p_1 values.

The simulation, analytical and simple model results for probability p_3 in Figure 5.7 show similar pattern as the results of p_1 and p_2. At low load of 100 devices, p_3 is below 0.1. But increasing the load results in higher values for p_3.

In case of probability p_4, the results for low load of 100 M2M devices show that it is highly unlikely that a 5th packet would arrive before the start of multiplexing process as depicted in Figure 5.8. However, the likeliness of a 5th arrival increases with increasing the M2M traffic load.

The results for probability p_5 in Figure 5.9 show that in the scenario of 100 M2M devices, it is not possible to multiplex 6 packets because the timer would

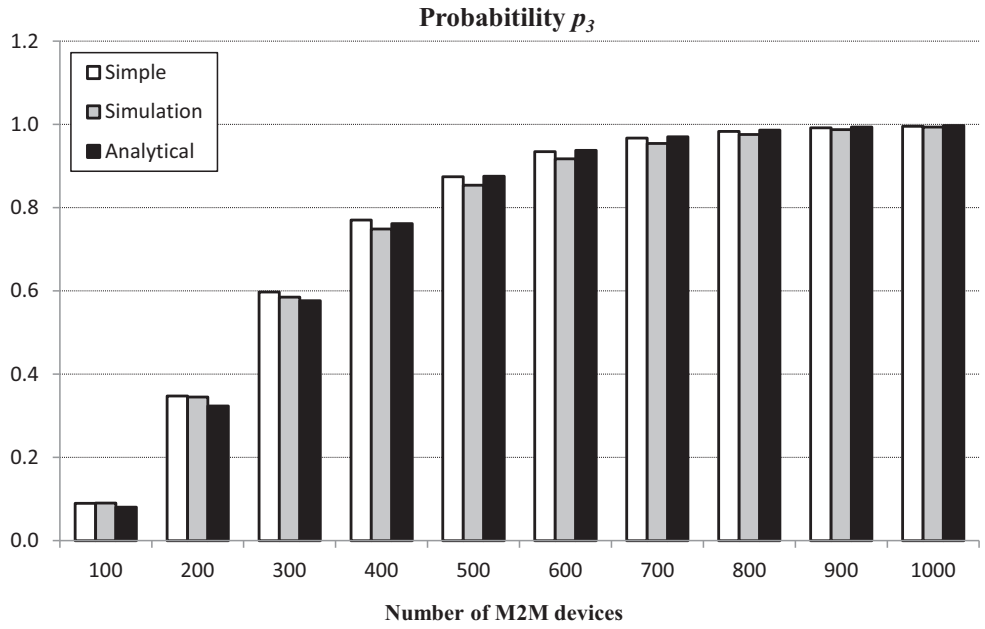

Figure 5.7: Comparison of probability p_3

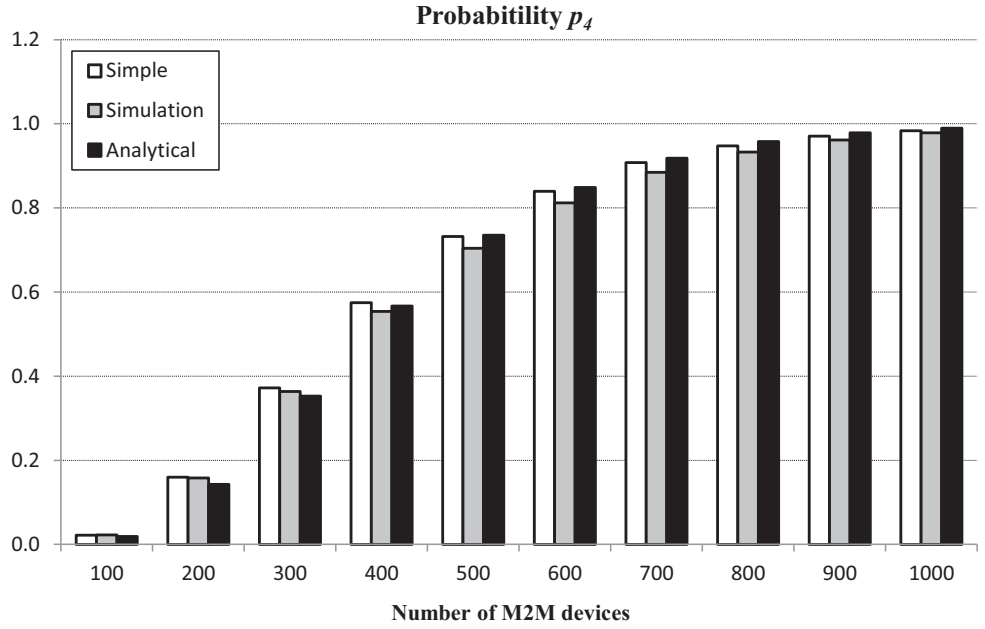

Figure 5.8: Comparison of probability p_4

not wait that long. Since the average interarrival time of packets is high, the timer plays a role every time packets are multiplexed. The arrival of r packets would not happen in this low load scenario. In higher load scenarios, the probability p_5 increases.

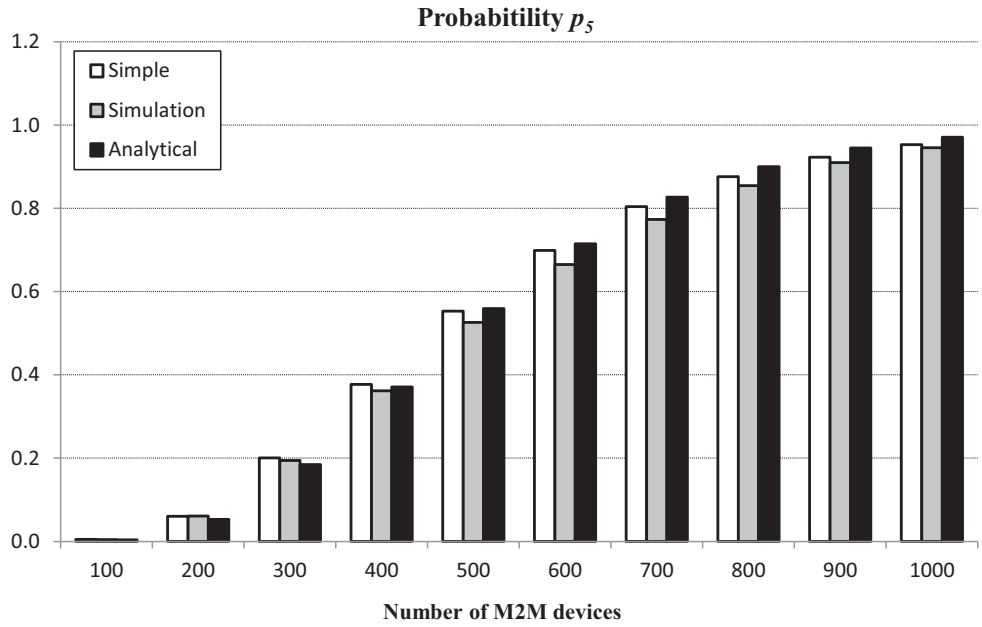

Figure 5.9: Comparison of probability p_5

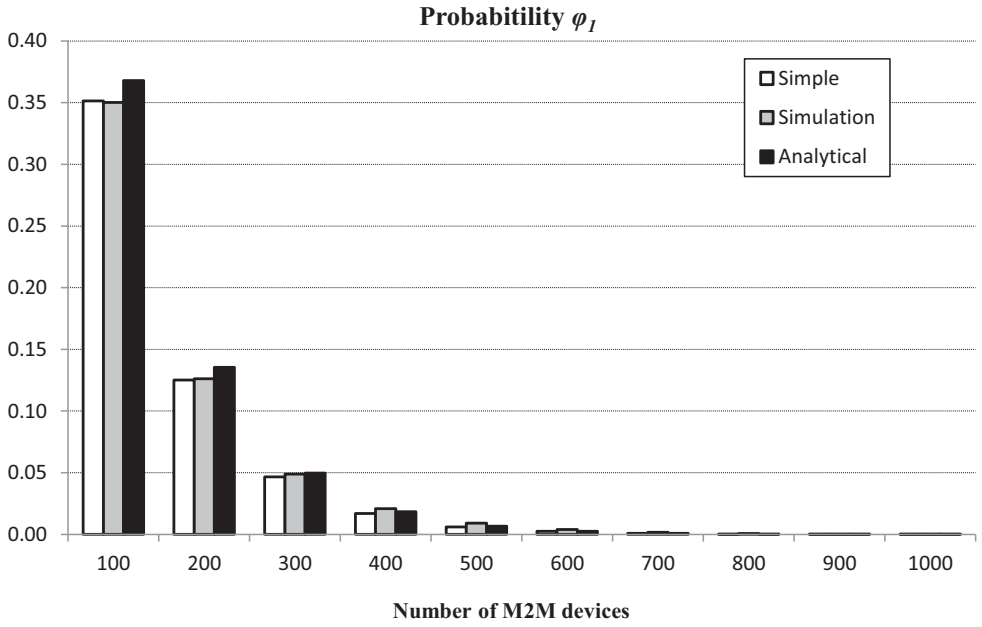

Figure 5.10: Comparison of probability ϕ_1

Values of path probabilities ϕ_k for different values of k are determined for low load scenarios. In case of 100 M2M users inside RN coverage area, the probability ϕ_1 is around 0.39 for simulation results and 0.37 for analytical and simple model results as in Figure 5.10. Increasing the traffic load decreases the probability ϕ_1 in subsequent scenarios. In case of 700 or more M2M devices in the RN coverage

area, ϕ_1 is very low.

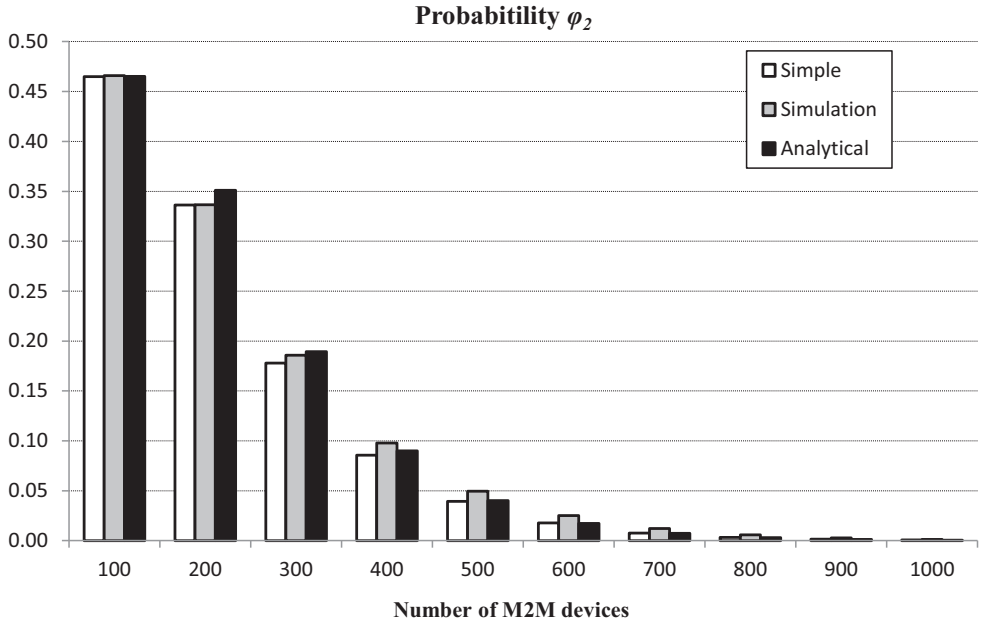

Figure 5.11: Comparison of probability ϕ_2

The results for ϕ_2 are shown in Figure 5.11. These results illustrate that the path probability ϕ_2 is high in low load scenarios as compared to high load scenarios (Figure 5.11).

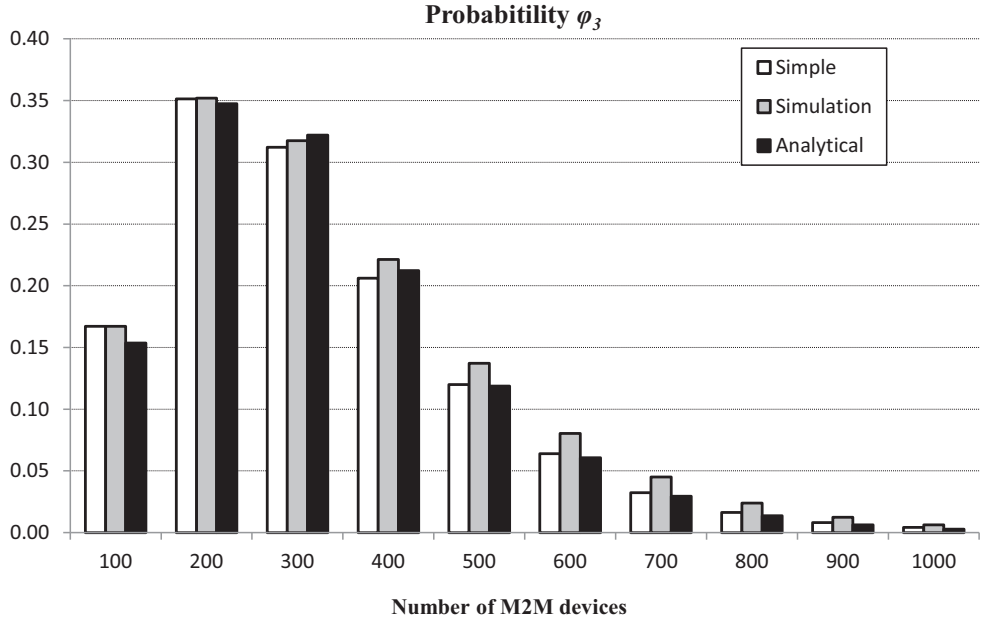

Figure 5.12: Comparison of probability ϕ_3

The results for probability ϕ_3 in Figure 5.12 show that the path probability ϕ_3 increases for load of 100 M2M devices to 200 M2M devices. However, increasing the load further results in gradual decrease in the probability ϕ_3.

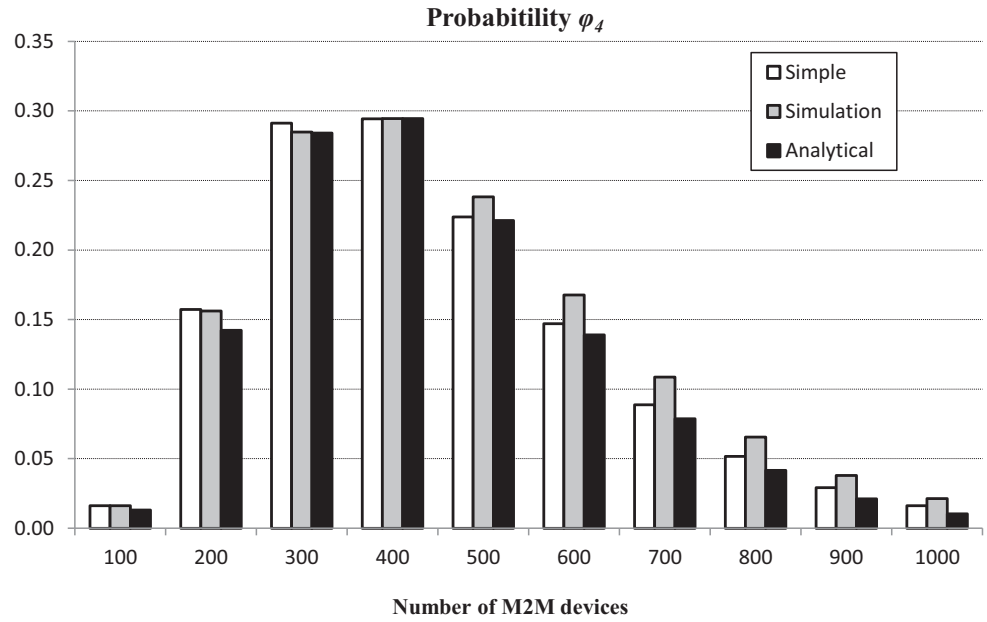

Figure 5.13: Comparison of probability ϕ_4

In case of probability ϕ_4, the results in Figure 5.13 are similar to results of ϕ_3. The probabilities increase initially when the load is increased to up to 400 devices. Further increase in load results in decrease in the probability ϕ_4.

The results for probability ϕ_5 in Figure 5.14 are similar to results in Figure 5.13.

5.4.2 Multiplexing Gain and Radio Resource Utilization

The process of aggregating data traffic at the RN and multiplexing packets of various M2M devices with timer ensures improved spectrum utilization. The simulation results in chapter 4 illustrated that the PRB usage significantly improves when the multiplexing scheme with timer is used at the RN as compared to relaying packets without multiplexing. In this section, the utilization of PRBs in each multiplexing process in various traffic load scenarios is derived analytically and the results of analytical model for PRBs usage are compared with simulation results for PRB usage. Furthermore, the results for the two cases, relaying without multiplexing as well as relaying with multiplexing and timer are also compared to obtain an insight of multiplexing gain achieved, both in simulation and analytical models.

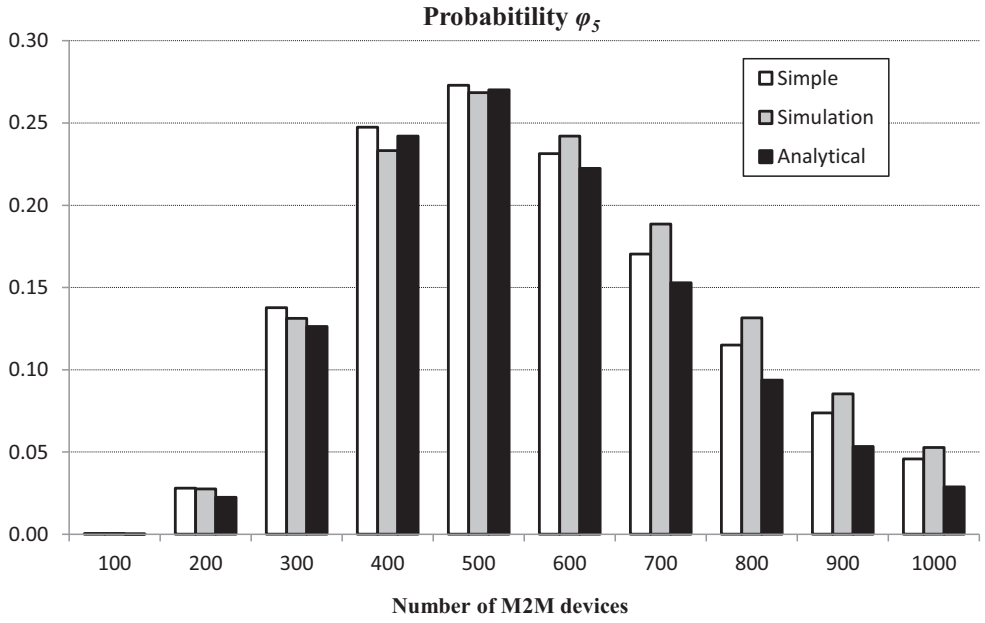

Figure 5.14: Comparison of probability ϕ_5

In order to increase the usage of PRBs in each multiplexing process, the timer functionality provides time for packets from different users to arrive at the RN and be multiplexed before transmission to DeNB. In case of traffic models where the packets arrive in bursts in between periods of few or no arrivals, the multiplexing process is quite beneficial. The traffic generation at the M2M devices is statistically independent. Therefore, the expected traffic at the RN would be in the form of bursts as well as periods of relative silence. So the primary objective of multiplexing is that instead of sending isolated packets in the form of small blocks (TBS) to DeNB, aggregate the arriving packets to form large blocks, reducing the total number of blocks sent to DeNB as a result. One advantage of this aggregation is that each block of data sent to DeNB requires 44 bytes (352 bits) share for overhead data. If the number of blocks is reduced and the block size is increased, considerable amount of data overhead can be avoided.

In order to determine the number of PRBs required for different stages of the r-stage coxian process (in other words, PRBs required for different buffer sizes), the TBS of different PRB allocations for the MCS has to be observed. The values for MCS 16 are given in Table 5.2.

When the r-stage coxian process enters stage 1, the buffer size can be up to l bits (29 bytes at the PDCP layer of RN Uu interface as in Table 5.1). Similarly, in stage k, the size can reach $k \times l$ bits. The number of PRBs required for each stage k of r-stage coxian process is $N_{k,MCS}$ which depends on stage k and MCS. The values

No. of PRBs	1	2	3	4	5
TBS (bytes)	41	79	121	161	201
TBS (bits)	328	632	968	1288	1608
Capacity excluding overhead (bytes)	0	70	77	117	157
Capacity excluding overhead (bits)	0	280	616	936	1256

Table 5.2: TBS capacity for various PRBs with MCS 16

of $N_{k,MCS}$ for MCS 16 and l 29 bytes (232 bits) are given in Table 5.3.

Packets in buffer	1	2	3	4	5	6
Buffer sizes (bytes)	29	58	87	116	145	174
Buffer sizes (bits)	232	464	696	928	1160	1392
PRBs required	2	3	4	4	5	5

Table 5.3: PRBs required for stages of r-stage coxian process with MCS 16 and l 232 bits

Let K be a discrete random variable representing the number of small packets multiplexed into a large packet. The possible values of K are $k = 1, 2, 3, ..., r$ where r is the maximum possible number of small packets in a large multiplexed packet. The probability mass function or pmf of K gives the probabilities of all possible values of K, denoted as $P[K = k]$.

Probability p_k is the likelihood of another arrival with k small packets already in the buffer before multiplexing. In other words, p_k is the probability that the value of K would be greater than k, i.e. the probability of having more than k small packets in a large multiplexed packet, which can be given as in (5.7)

$$P[K > k] = 1 - \left(\sum_{j=0}^{k-1} \frac{e^{-\lambda T_{\max}} (\lambda T_{\max})^j}{j!} \right) \text{ for } 1 \leq k < r \tag{5.7}$$

For example, if $p_1 = 0.6$, it means that 60% of the large multiplexed packets have size greater than 1, i.e. $P[K > 1] = 0.6$. The probability $P[K > 0]$ (k equal to 0) should be equal to 1 because multiplexing always require at least one small packet, i.e. all values of K are greater than 0. Putting $k = 0$ in (5.7) gives $P[K > 0] = 1$. Similarly, the probability $P[K \geq r]$ (k greater than or equal to r) should be 0. Since, K is a discrete random variable, the probability $P[K = k]$ can be determined as

$$P[K = k] = P[K > k - 1] - P[K > k] \tag{5.8}$$

Equation (5.7) gives

$$\begin{aligned} P[K=k] &= \left(1-\sum_{j=0}^{k-2}\frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^j}{j!}\right) \\ &\quad -\left(1-\sum_{j=0}^{k-1}\frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^j}{j!}\right) \end{aligned} \tag{5.9}$$

$$\Rightarrow P[K=k] = -\sum_{j=0}^{k-2}\frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^j}{j!} + \sum_{j=0}^{k-1}\frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^j}{j!} \tag{5.10}$$

$$\begin{aligned} \Rightarrow P[K=k] &= -\sum_{j=0}^{k-2}\frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^j}{j!} \\ &\quad +\sum_{j=0}^{k-2}\frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^j}{j!} + \frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^{k-1}}{(k-1)!} \end{aligned} \tag{5.11}$$

$$\Rightarrow P[K=k] = \frac{e^{-\lambda T_{\max}}(\lambda T_{\max})^{k-1}}{(k-1)!} \tag{5.12}$$

From $N_{k,MCS}$ and $P[K=k]$, the mean number of PRBs used in each multiplexing process in a particular traffic load scenario, $\overline{N}$ can be determined as

$$\overline{N} = \sum_{k=1}^{r} N_{k,MCS} \times P[K=k] \tag{5.13}$$

$\overline{N}$ is the average number of PRBs used by one large multiplexed packet. It is different from the average PRBs used per TTI. The average number of PRB usage per TTI is significantly reduced by using the multiplexing with timer scheme as shown in Figure 4.15 and Figure 4.18. In case of $\overline{N}$, the goal is to determine the increase in PRB usage for each multiplexing process. In the multiplexing buffer, the multiplexing process occurs after a period of aggregation of small packets in case of multiplexing with timer. The duration of this period is defined by the parameter $T_{\max}$. In case of no multiplexing, the arriving packets are sent ahead immediately at the beginning of the next TTI. Therefore, for the case of multiplexing with timer, the average PRB usage for each large IP packet would increase and the overall number of large packets to be sent to DeNB would be significantly reduced. This would also result in significantly reduced overhead data traffic.

Based on the definition of $\overline{N}$ in (5.13), the simulation and analytical results are compared for the cases of relaying without multiplexing as well as relaying with multiplexing and timer in Figure 5.15. The 95% confidence intervals for simulation results are provided in Appendix.

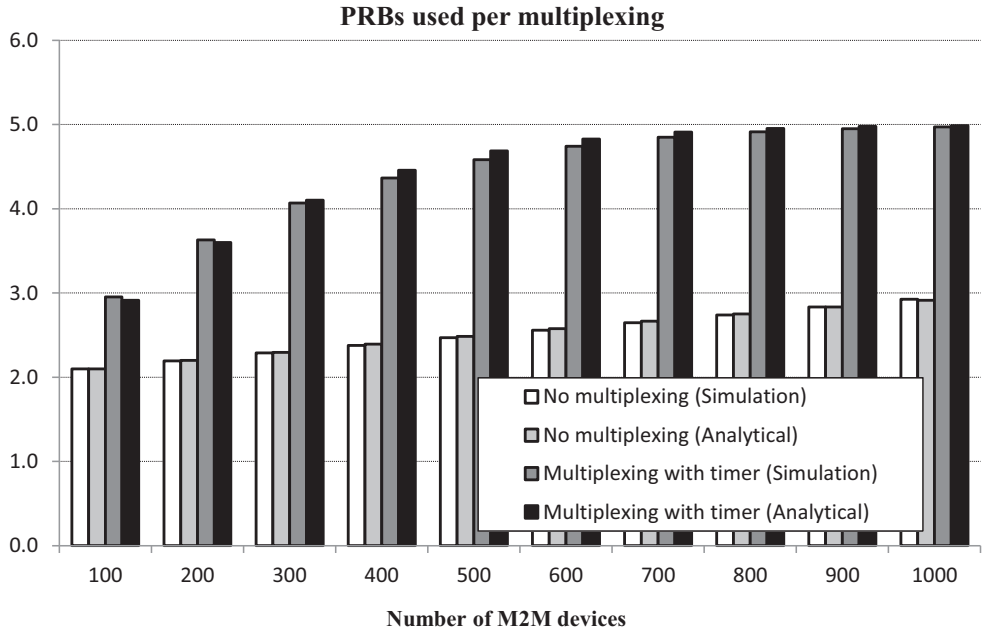

Figure 5.15: Comparison of $\overline{N}$ with and without multiplexing

There are two aspects of Figure 5.15 to highlight. Firstly, the comparison of results for relaying with multiplexing and without multiplexing show that the radio resource utilization improves significantly by avoiding excessive overhead data. In case of no multiplexing, the average usage of PRB per large IP packet is low, which indicates that the number of blocks transmitted from RN to DeNB is quite high. As a result, the overhead data also grows. The graph depicting results for multiplexing with timer shows that packets are multiplexed into larger blocks. This helps in reducing the overhead data required for each transmission block. Secondly, the simulation results are in accordance with analytical results. The agreement between simulation and analytical results helps in validation of both models. The multiplexing gain achieved in each scenario is illustrated in Figure 5.16. In order to compute the multiplexing gain in each scenario, the $\overline{N}$ for that particular scenario without multiplexing (denoted as $\overline{N}_{no\_mux}$) and with multiplexing and timer (denoted as $\overline{N}_{mux\_t}$) from (5.13) is used as in (5.14)

$$gain\% = 100 - \left(\frac{\overline{N}_{no\_mux}}{\overline{N}_{mux\_t}} \times 100 \right) \tag{5.14}$$

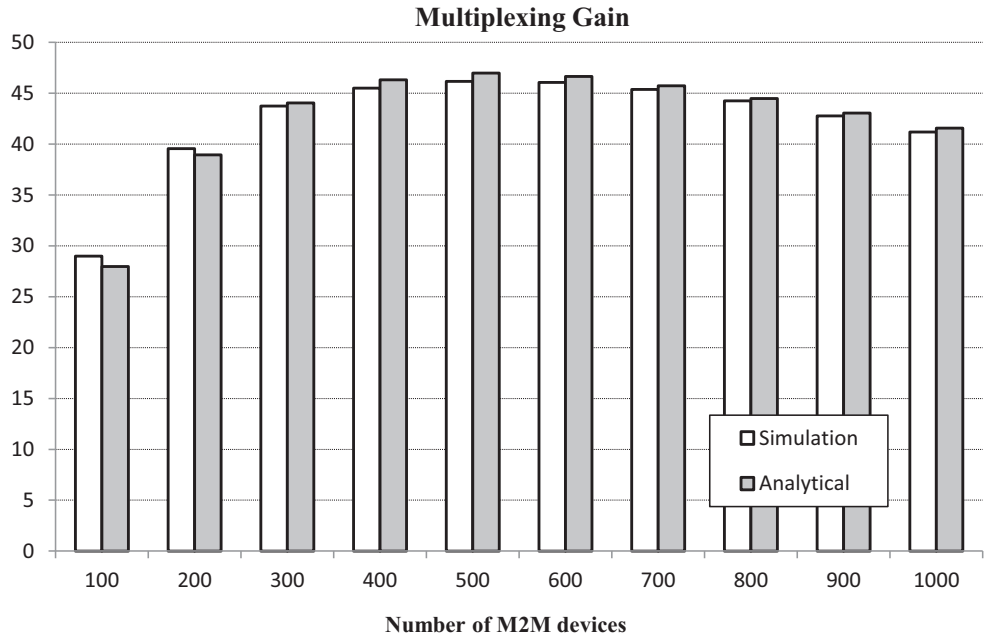

Figure 5.16: Multiplexing gain comparison in simulation and analytical models

5.5 Summary

The goal of analytical model was to validate the results obtained through simulations using a mathematical approach. The analytical model presented in this chapter provided results for p_k, ϕ_k, $\overline{N}$ and multiplexing gain. This model provides a base for developing analytical models for transmission delays for traffic of various QoS classes.

6 Conclusion and Outlook

6.1 Conclusion

This thesis focused on 3GPP based mobile networks and future M2M communication. The primary goal of this thesis was to investigate methods for incorporating future M2M data traffic into LTE-A networks. In order to achieve this, a relaying based concept was developed for integrating the M2M communication into LTE-A. The features and functionalities of LTE and LTE-A related to air interface were highlighted to provide a background for the problem statement. A literature survey was presented to provide an overview of the related research work done recently.

Radio resource allocation is a vital functionality in LTE and LTE-A networks for efficient spectrum utilization. A radio resource allocation scheme for SC-FDMA based LTE uplink was reviewed and evaluated in terms of fairness and QoS performance. The environment for the design, implementation and analysis of the radio scheduler was developed using the OPNET Modeler simulator. Channel models, mobility models, LTE architecture and LTE protocol stack were developed for facilitating the simulation environment. The model was further extended to incorporate the LTE-A air interface scheme and Carrier Aggregation (CA) feature. A Component Carrier Selection (CCS) scheme was presented and evaluated by comparing it with other reference schemes. The design of an LTE-A uplink scheduler was also discussed and evaluated.

An overview of the M2M communication was provided by discussing the architecture, standardization and application areas of M2M communication. The issues related to future M2M traffic such as mass deployment of M2M devices in the near future were elaborated. The impact of M2M data traffic on the QoS performance of regular LTE and LTE-A traffic such as VoIP and video was investigated. It was shown that the future M2M data traffic is capable of posing capacity and performance challenges. The M2M traffic has usually narrowband requirements, whereas the LTE-A has been designed for broadband data service.

The RN was proposed as a solution to the problem of M2M traffic integration into LTE-A networks. RN offers low cost and low power solution to issues like coverage extension and cell edge throughput enhancement. An RN based frame-

work was presented for aggregating and multiplexing the M2M data at the RN before transmission over the air interface. As a result, it was illustrated that valuable and scarce radio resources could be efficiently utilized. The RN also minimized the impact of M2M traffic on the QoS performance of the regular LTE-A traffic. The improvement in the cell edge user performance was also shown. The simulation environment was designed according to relaying protocol stack.

The designed framework is capable of simulating RNs in a hexagonal regular grid setup with the center cell being surrounded by another tier of six cells. Any number of RNs can be connected to any eNodeB. The parameters like the maximum number of PRBs for an RN $N_{\max}$, the location of RN, the timer expiry duration $T_{\max}$ etc. can be tuned to modify the behavior of the network.

The model is formulated mathematically as well in the form of an analytical model for comparison with simulation results. The multiplexing transition probabilities, PRB utilization and multiplexing gain are analytically modeled and found to be in agreement with simulation results.

6.2 Outlook

The proposed framework has been specifically designed for M2M data aggregation and multiplexing. However, the framework can be further extended for the aggregation and multiplexing of VoIP packets as well. Furthermore, the design of DeNB can be modified to introduce aggregation and multiplexing for downlink data traffic.

The RN scheduler can be improved further by introducing QoS differentiation among the devices and UEs served by the RN. The M2M applications such as Intelligent Traffic System (ITS), e-healthcare etc. might have delay sensitive data traffic. Smart metering data traffic can be considered as best effort traffic. This differentiation can be extremely helpful in mixed traffic scenarios with both M2M and regular LTE-A traffic served by the RN.

The RN design can also be extended from fixed RN to moving RN with varying channel conditions. This can be a good starting point for investigating several new M2M application areas and determining the feasibility of incorporating them into LTE-A networks.

The analytical model can also be extended to obtain results for multiplexing delay mathematically. Similarly, the QoS performance improvement and transmission delays can also be analytically derived.

7 List of Publications

7.1 Journal Papers

1. S. N. K. Marwat, T. Weerawardane, Y. Zaki, C. Goerg and A. Timm-Giel, "Analysis of Radio Resource Allocation in LTE Uplink", *Wireless Personal Communications*, Volume 79, Issue 3, Dec. 2014, pp. 2305-2322, ISSN 0929-6212, Springer, New York, USA.

2. S. N. K. Marwat, S. Meyer, T. Weerawardane, and C. Goerg, "Congestion Aware Handover in LTE System for Load Balancing in Transport Network", *ETRI Journal*, Volume 36, Issue 5, Oct. 2014, pp. 761-771, ISSN 1225-6463, Electronics Telecommunications Research Institute, Daejeon, Republic of Korea.

3. S. N. K. Marwat, T. Weerawardane, Y. Zaki, and C. Goerg, "Radio Resource Allocation for M2M Communication in LTE Uplink Scheduling", *Rügge, I. and Himstedt, A. (eds.): LogDynamics Research Report*; Volume 3, 2012/13, pp. 31-34, ISSN 1867-0210, Bremen Research Cluster for Dynamics in Logistics, Universität Bremen.

7.2 Conference Papers

1. Y. Mehmood, S. N. K. Marwat, C. Goerg, Y. Zaki, and A. Timm-Giel, "Evaluation of M2M Data Traffic Aggregation in LTE-A Uplink", *20. VDE/ITG Fachtagung Mobilkommunikation*, Osnabrück, Germany on 7 May 2015, pp. 51-55.

2. S. N. K. Marwat, Yangyang Dong, Xi Li, Y. Zaki and C. Goerg, "Novel Schemes for Component Carrier Selection and Radio Resource Allocation in LTE-Advanced Uplink", *6th International Conference on Mobile Networks and Management*, Wuerzburg, Germany, on 22-24 Sept. 2014.

3. F. Ahmad, S. N. K. Marwat, Y. Zaki, and C. Goerg, “Tailoring LTE-Advanced for M2M Communication using Wireless Inband Relay Node”, *World Telecommunications Congress*, Berlin, Germany, on 1-3 June 2014.

4. Y. Mehmood, T. Pötsch, S. N. K. Marwat, and C. Goerg, “Impact of M2M Traffic on LTE Data Traffic Performance”, *4th International Conference on Dynamics in Logistics*, Bremen, Germany, 10-14 February 2014.

5. F. Ahmad, S. N. K. Marwat, Y. Zaki, and C. Goerg, “Machine-to-Machine Sensor Data Multiplexing using LTE-Advanced Relay Node for Logistics”, *4th International Conference on Dynamics in Logistics*, Bremen, Germany, 10-14 February 2014.

6. S. N. K. Marwat, Y. Zaki, J. Chen, A. Timm-Giel, and C. Goerg, “A Novel Machine-to-Machine Traffic Multiplexing in LTE-A System using Wireless In-band Relaying”, *5th International Conference on Mobile Networks and Management*, Cork, Ireland, 23-25 September 2013.

7. H. Zhang, N. El-Berishy, D. Zastrau, S. N. K. Marwat, Y. Tan, J. Palafox, and I. Rügge, “Interdisciplinary Perspectives of Knowledge Management in Logistics”, *Logistikmanagement*, Bremen, Germany, 9-13 September 2013.

8. S. N. K. Marwat, T. Pötsch, Y. Zaki, T. Weerawardane, and C. Goerg, “Addressing the Challenges of E-Healthcare in Future Mobile Networks”, *19th EUNICE Workshop on Advances in Communication Networking*, Chemnitz, Germany, 28-30 August 2013, pp. 90-99.

9. T. Pötsch, S. N. K. Marwat, Y. Zaki, C. Goerg, “Influence of Future M2M Communication on the LTE system”, *Wireless and Mobile Networking Conference*, Dubai, United Arab Emirates on 23-25 April, 2013.

10. D. Awan, S. Bashir, S. N. K. Marwat and H. Ali, “Gain enhancement of UWB antennas with and without band notch feature”, *International Workshop on Antenna Technology*, Karlsruhe, Germany on 4-6 March 2013, pp. 342-345.

11. S. N. K. Marwat, T. Weerawardane, Y. Zaki, C. Goerg, and A. Timm-Giel, “Design and Performance Analysis of Bandwidth and QoS Aware LTE Uplink Scheduler in Heterogeneous Traffic Environment”, *The 8th International Wireless Communications and Mobile Computing Conference*, Limassol, Cyprus on 27-31 August 2012, pp. 499-504.

12. S. N. K. Marwat, T. Weerawardane, R. Perera, C. Goerg, and A. Timm-Giel, "Impact of Machine-to-Machine (M2M) Communications on Disaster Management in Future Mobile Networks", *The General Sir John Kotelawala Defense University Annual Symposium*, Ratmalana, Sri Lanka, 22-23 August 2012.

13. S. N. K. Marwat, T. Weerawardane, Y. Zaki, C. Goerg, and A. Timm-Giel, "Performance Evaluation of Bandwidth and QoS Aware LTE Uplink Scheduler", *10th International Conference on Wired/Wireless Internet Communications*, Santorini, Greece, on 6-8 June 2012, pp. 298-306.

14. S. N. K. Marwat, T. Weerawardane, Y. Zaki, C. Goerg, and A. Timm-Giel, "Performance of Bandwidth and QoS Aware LTE Uplink Scheduler Towards Delay Sensitive Traffic", *17. VDE/ITG Fachtagung Mobilkommunikation*, Osnabrück, Germany on 9-10 May 2012, pp. 51-55.

7.3 Posters

1. T. Pötsch, S. N. K. Marwat, Y. Zaki, and C. Goerg, "Performance Evaluation of Machine-to-Machine Communication on Future Mobile Networks in Disaster Scenarios", *19th EUNICE Workshop on Advances in Communication Networking*, Chemnitz, Germany on 28-30 August 2013.

2. S. N. K. Marwat and C. Goerg, "Machine-to-Machine Communications for eHealthcare in Future Mobile Networks", *ICT meets Medicine and Health - Joint CEWIT-TZI-acatech Workshop*, Bremen, Germany on 19-20 March 2013.

Appendix

A Appendix Chapter

A.1 Confidence Intervals for Simulation Results

The 95% confidence intervals for simulation results in chapter 4 are provided in Table A.1 to Table A.14.

Number of video users	**10**	**20**	**30**	**40**	**50**	**60**
Without relaying	0.018	0.023	0.024	0.021	0.008	0
With relaying	0.018	0.021	0.024	0.022	0.018	0

Table A.1: Confidence intervals for Figure 4.11

Number of video users	**10**	**20**	**30**	**40**	**50**	**60**
Without relaying	9.12	9.14	8.83	9.55	25.60	0
With relaying	8.72	8.70	8.94	10.90	9.03	0

Table A.2: Confidence intervals divided by 10^{-6} for Figure 4.12

Number of video users	**10**	**20**	**30**	**40**	**50**	**60**
Without relaying	5.25	6.46	6.86	5.97	2.39	0
With relaying	5.66	6.86	7.61	6.97	5.63	0

Table A.3: Confidence intervals for Figure 4.13

Number of FTP users	**5**	**10**	**15**	**20**	**25**
Without relaying	0.043	0.155	0.168	2.001	0.553
With relaying	0.049	0.099	0.310	2.328	3

Table A.4: Confidence intervals for Figure 4.14

Number of M2M devices	**1000**	**2000**	**3000**	**4000**	**5000**
No multiplexing	0.0040	0.0053	0.0056	0.0055	0.0053
Multiplexing without timer	0.0016	0.0021	0.0025	0.0029	0.0032
Multiplexing with timer	0.0014	0.002	0.0025	0.0029	0.0032
Number of M2M devices	**6000**	**7000**	**8000**	**9000**	**10000**
No multiplexing	0.0049	0	0	0	0
Multiplexing without timer	0.0035	0.0038	0.0040	0.0042	0.0048
Multiplexing with timer	0.0035	0.0038	0.0040	0.0042	0.0048
Number of M2M devices	**11000**	**12000**	**13000**	**14000**	**15000**
No multiplexing	0	0	0	0	0
Multiplexing without timer	0.0046	0.0049	0.0051	0.0047	0
Multiplexing with timer	0.0046	0.0049	0.0051	0.0047	0

Table A.5: Confidence intervals for Figure 4.15

Number of M2M devices	**1000**	**2000**	**3000**	**4000**	**5000**
No multiplexing	1.20	0.88	0.88	1.16	3.54
Multiplexing without timer	37.65	13.34	7.55	5.05	3.79
Multiplexing with timer	13.27	11.20	7.38	5.11	3.83
Number of M2M devices	**6000**	**7000**	**8000**	**9000**	**10000**
No multiplexing	199.5143	-	-	-	-
Multiplexing without timer	3.10	2.64	2.26	2.01	2
Multiplexing with timer	3.08	2.61	2.26	2.01	2
Number of M2M devices	**11000**	**12000**	**13000**	**14000**	**15000**
No multiplexing	-	-	-	-	-
Multiplexing without timer	3.63	53.23	150.80	300.28	-
Multiplexing with timer	3.63	53.23	150.80	300.28	-

Table A.6: Confidence intervals divided by 10^{-6} for Figure 4.16

Number of M2M devices	**1000**	**2000**	**3000**	**4000**	**5000**
No multiplexing	2.29	3.18	3.77	4.03	3.86
Multiplexing without timer	3.85	5.24	6.15	6.79	7.23
Multiplexing with timer	3.34	5.16	6.15	6.79	7.23
Number of M2M devices	**6000**	**7000**	**8000**	**9000**	**10000**
No multiplexing	2.98	0	0	0	0
Multiplexing without timer	7.50	7.62	7.59	7.42	7.66
Multiplexing with timer	7.50	7.62	7.59	7.42	7.66
Number of M2M devices	**11000**	**12000**	**13000**	**14000**	**15000**
No multiplexing	0	0	0	0	0
Multiplexing without timer	6.60	5.87	4.83	3.19	0
Multiplexing with timer	6.60	5.87	4.83	3.19	0

Table A.7: Confidence intervals for Figure 4.17

Number of M2M devices	**1000**	**2000**	**3000**	**4000**	**5000**	**6000**
No multiplexing	0.0075	0.0096	0	0	0	0
Multiplexing without timer	0.0038	0.0054	0.0065	0.0074	0.0074	0
Multiplexing with timer	0.0037	0.0054	0.0066	0.0075	0.0074	0

Table A.8: Confidence intervals for Figure 4.18

Number of M2M devices	**1000**	**2000**	**3000**	**4000**	**5000**	**6000**
No multiplexing	2.06	5.71	-	-	-	-
Multiplexing without timer	15.18	5.57	3.17	3.46	293.93	-
Multiplexing with timer	12.93	5.52	3.20	3.46	293.93	-

Table A.9: Confidence intervals divided by 10^{-6} for Figure 4.19

Number of M2M devices	**1000**	**2000**	**3000**	**4000**	**5000**	**6000**
No multiplexing	2.13	2.36	0	0	0	0
Multiplexing without timer	2.54	3.15	3.25	2.88	1.80	0
Multiplexing with timer	2.52	3.15	3.25	2.87	1.80	0

Table A.10: Confidence intervals for Figure 4.20

Number of FTP users	**1000**	**2000**	**3000**	**4000**	**5000**
Without relaying	0	0	0	0	0
With relaying	0	0	0	0	0

Table A.11: Confidence intervals for Figure 4.21

Number of FTP users	1000	2000	3000	4000	5000
Without relaying	3.737	2.708	3.201	3.069	-
With relaying	14.915	9.038	5.360	3.748	2.805

Table A.12: Confidence intervals divided by 10^{-6} for Figure 4.22

Number of FTP users	1000	2000	3000	4000	5000
Without relaying	0.141	0.150	0.163	0.583	1.366
With relaying	0.127	0.133	0.147	0.141	0.152

Table A.13: Confidence intervals for Figure 4.23

Number of FTP users	1000	2000	3000	4000	5000
Without relaying	0.761	1.048	0.976	1.140	1.093
With relaying	0.413	0.686	0.790	0.845	0.862

Table A.14: Confidence intervals for Figure 4.24

A.2 Confidence Intervals for Comparison of Simulation and Analytical Results

The 95% confidence intervals for simulation results in chapter 5 are provided in Table A.15 to Table A.24.

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	1393.01	555.01	477.10	361.54	136.75
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	165.14	57.38	43.80	31.63	19.68

Table A.15: Confidence intervals divided by 10^{-6} for Figure 5.5

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	1605.34	914.96	1225.19	1312.07	413.37
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	232.01	163.25	134.31	86.45	60.64

Table A.16: Confidence intervals divided by 10^{-6} for Figure 5.6

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	1013.09	816.60	1240.36	1311.17	807.50
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	520.58	509.94	489.09	96.70	136.70

Table A.17: Confidence intervals divided by 10^{-6} for Figure 5.7

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	613.20	676.92	846.83	1244.80	969.50
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	810.40	727.65	496.34	292.33	288.41

Table A.18: Confidence intervals divided by 10^{-6} for Figure 5.8

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	227.88	486.46	753.57	1108.50	1072.97
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	758.09	404.01	648.26	406.56	214.17

Table A.19: Confidence intervals divided by 10^{-6} for Figure 5.9

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	1393.01	555.01	477.10	361.54	136.75
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	165.14	57.38	43.80	31.63	19.68

Table A.20: Confidence intervals divided by 10^{-6} for Figure 5.10

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	1140.06	706.25	1158.67	1260.59	406.49
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	230.58	162.40	134.09	86.47	60.64

Table A.21: Confidence intervals divided by 10^{-6} for Figure 5.11

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	1017.31	845.81	551.77	964.85	731.13
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	496.34	496.90	483.58	96.61	136.14

Table A.22: Confidence intervals divided by 10^{-6} for Figure 5.12

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	260.21	438.24	1059.16	756.47	618.12
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	665.69	632.93	459.38	286.65	283.72

Table A.23: Confidence intervals divided by 10^{-6} for Figure 5.13

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Simulation	15.00	152.32	517.05	848.27	375.72
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Simulation	701.41	191.14	512.61	361.46	190.76

Table A.24: Confidence intervals divided by 10^{-6} for Figure 5.14

Number of M2M devices	**100**	**200**	**300**	**400**	**500**
No multiplexing	0.0012	0.0007	0.0008	0.0012	0.001
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
No multiplexing	0.0017	0.00127	0.00097	0.00077	0.0011
Number of M2M devices	**100**	**200**	**300**	**400**	**500**
Multiplexing with timer	0.0031	0.0026	0.002	0.0015	0.0022
Number of M2M devices	**600**	**700**	**800**	**900**	**1000**
Multiplexing with timer	0.0011	0.0008	0.0007	0.0005	0.0003

Table A.25: Confidence intervals for Figure 5.15

A.3 3GPP Transport Block Size Table

This is a subset of the full table from [36.10a]. Each combination of an MCS and number of PRBs results in a TBS (in bits).

	Number of PRBs									
MCS	1	2	3	4	5	6	7	8	9	10
0	16	32	56	88	120	152	176	208	224	256
1	24	56	88	144	176	208	224	256	328	344
2	32	72	144	176	208	256	296	328	376	424
3	40	104	176	208	256	328	392	440	504	568
4	56	120	208	256	328	408	488	552	632	696
5	72	144	224	328	424	504	600	680	776	872
6	328	176	256	392	504	600	712	808	936	1032
7	104	224	328	472	584	712	840	968	1096	1224
8	120	256	392	536	680	808	968	1096	1256	1384
9	136	296	456	616	776	936	1096	1256	1416	1544
10	144	328	504	680	872	1032	1224	1384	1544	1736
11	176	376	584	776	1000	1192	1384	1608	1800	2024
12	208	440	680	904	1128	1352	1608	1800	2024	2280
13	224	488	744	1000	1256	1544	1800	2024	2280	2536
14	256	552	840	1128	1416	1736	1992	2280	2600	2856
15	280	600	904	1224	1544	1800	2152	2472	2728	3112
16	328	632	968	1288	1608	1928	2280	2600	2984	3240
17	336	696	1064	1416	1800	2152	2536	2856	3240	3624
18	376	776	1160	1544	1992	2344	2792	3112	3624	4008
19	408	840	1288	1736	2152	2600	2984	3496	3880	4264
20	440	904	1384	1864	2344	2792	3240	3752	4136	4584
21	488	1000	1480	1992	2472	2984	3496	4008	4584	4968
22	520	1064	1608	2152	2664	3240	3752	4264	4776	5352
23	552	1128	1736	2280	2856	3496	4008	4584	5160	5736
24	584	1192	1800	2408	2984	3624	4264	4968	5544	5992
25	616	1256	1864	2536	3112	3752	4392	5160	5736	6200
26	712	1480	2216	2984	3752	4392	5160	5992	6712	7480

Table A.26: 3GPP table of TBS

Bibliography

[10198] ETSI Technical Report TR 101.112. Selection procedures for the choice of radio transmission technologies of the UMTS. Technical Report version 3.2.0, European Telecommunications Standards Institute, Apr. 1998.

[10213] ETSI Technical Specification TS 102.690. Machine-to-Machine communications (M2M); Functional architecture. Technical Report version 2.1.1, European Telecommunications Standards Institute, Oct. 2013.

[23.12] 3GPP Technical Report TR 23.888. Technical Specification Group Services and System Aspects; System improvements for Machine-Type Communications. Technical Report version 11.0.0, 3rd Generation Partnership Project, Sept. 2012.

[25.09] 3GPP Technical Report TR 25.913. Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN). Technical Report version 9.0.0, 3rd Generation Partnership Project, Dec. 2009.

[33.11] 3GPP Technical Report TR 33.816. Technical Specification Group Services and System Aspects; Feasibility study on LTE relay node security. Technical Report version 10.0.0, 3rd Generation Partnership Project, Mar. 2011.

[36.10a] 3GPP Technical Specifications TS 36.213. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures. Technical Report version 9.2.0, 3rd Generation Partnership Project, June 2010.

[36.10b] 3GPP Technical Report TR 36.814. Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects. Technical Report version 9.0.0, 3rd Generation Partnership Project, Mar. 2010.

[36.12] 3GPP Technical Report TR 36.912. Feasibility study for Further Advancements for E-UTRA (LTE-Advanced). Technical Report version 11.0.0, 3rd Generation Partnership Project, Sept. 2012.

[36.13] 3GPP Technical Specification TS 36.300. Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2. Technical Report Version 11.6.0, 3rd Generation Partnership Project, June 2013.

[ARC+08] M. Anas, C. Rosa, F. D. Calabrese, K. I. Pedersen, and P. E. Mogensen. Combined Admission Control and Scheduling for QoS Differentiation in LTE Uplink. In *IEEE 68th Vehicular Technology Conference*, Calgary, Alberta, Canada, 21-24 Sept. 2008.

[AY11] T. Ali-Yahiya. *Understanding LTE and Its Performance*. Springer, ISBN 9781441964564, LCCN 2011929037, 2011.

[Bal10] Balamurali. Optimal Downlink Control Channel Resource Allocation for LTE Systems. In *International Conference on Signal Processing and Communications*, Bangalore, India, 18-21 July 2010.

[BL12] G. J. Bradford and J. N. Laneman. Low Latency Relaying Schemes for Next-Generation Cellular Networks. In *IEEE International Conference on Communications*, pages 4294–4299, Ottawa, ON, Canada, 10-15 June 2012.

[BMS12] A. Baid, R. Madan, and A. Sampath. Delay Estimation and Fast Iterative Scheduling Policies for LTE Uplink. In *10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pages 89–96, Paderborn, Germany, 14-18 May 2012.

[BQC+08] M. Boussif, N. Quintero, F. D. Calabrese, C. Rosa, and J. Wigard. Interference Based Power Control Performance in LTE Uplink. In *IEEE International Symposium on Wireless Communication Systems*, pages 698–702, Reykjavík, Iceland, 21-24 Oct. 2008.

[BRRH10] O. Bulakci, S. Redana, B. Raaf, and J. Hamalainen. Performance Enhancement in LTE-Advanced Relay Networks via Relay Site Planning. In *IEEE 71st Vehicular Technology Conference*, Taipei, Taiwan, 16-19 May 2010.

[Bul10] Ö Bulakci. Multi-hop Moving Relays for IMT-Advanced and Beyond. In *Licentiate Seminar*, Department of Communications and Networking, Aalto University, Finland, 2010.

[Cal09] F. D. Calabrese. Scheduling and Link Adaptation for Uplink SC-FDMA Systems. In *PhD Thesis*. the Faculty of Engineering, Science and Medicine, Aalborg University, Aalborg, Denkmark, Apr. 2009.

[Cav00] J. K. Cavers. *Mobile Channel Characteristics*. Kluwer Academic Publishers, ISBN 9780792379263, LCCN 00056134, 2000.

[CE11] L. Coetzee and J. Eksteen. The Internet of Things - Promise for the Future? An Introduction. In *IST-Africa Conference Proceedings*, Gaborone, Botswana, 11-13 May 2011.

[CG14] R. Cohen and G. Grebla. Multidimensional OFDMA Scheduling in a Wireless Network with Relay Nodes. *IEEE/ACM Transactions on Networking*, PP(99), Association for Computing Machinery, ISSN 1063-6692, 14 Aug. 2014.

[Cho11] HanGyu Cho. Machine to Machine (M2M) Communications Technical Report. Technical report, IEEE 802.16 Broadband Wireless Access Working Group, 11 Nov. 2011.

[Cis14] Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013-2018. Technical Report Digital Publication, Cisco Systems Inc., 5 Feb. 2014.

[Cox12] C. Cox. *An Introduction to LTE: LTE, LTE-Advanced, SAE and 4G Mobile Communications*. John Wiley and Sons, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, ISBN 9781119943532, LCCN 2011047216, 2012.

[CTZC02] A. F. Canton, S. Tohme, D. Zeghlache, and T. Chahed. Performance Analysis of AAL2/ATM in UMTS Radio Access Network. In *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1352–1356 vol.3, Lisbon, Portugal, 15-18 Sept. 2002.

[CWL12] Min Chen, Jiafu Wan, and Fang Li. Machine-to-Machine Communications: Architectures, Standards and Applications. *KSII Transactions on Internet and Information Systems*, 6(2):480–497, Korean Society for Internet Information, ISSN 1976-7277, Feb. 2012.

[CYG95] Chen Nee Chuah, R. D. Yates, and D. J. Goodman. Integrated Dynamic Radio Resource Management. In *IEEE 45th Vehicular Technology Conference*, pages 584–588 vol.2, Chicago, IL, USA, 25-28 July 1995.

[dMBS13] T. M. de Moraes, G. Bauch, and E. Seidel. QoS-aware Scheduling for In-Band Relays in LTE-Advanced. In *Proceedings of 9th International ITG Conference on Systems, Communication and Coding*, Munich, Germany, 21-24 Jan. 2013.

[dMSBS13] T. M. de Moraes, A. B. Saleh, G. Bauch, and E. Seidel. QoS-Aware Traffic Scheduling in LTE-Advanced Relay-Enhanced Networks. In *IEEE 77th Vehicular Technology Conference*, Dresden, Germany, 2-5 June 2013.

[DPS11] E. Dahlman, S. Parkvall, and J. Skold. *4G LTE/LTE-Advanced for Mobile Broadband*. Academic Press, ISBN 9780123854896, LCCN 2011921244, 2011.

[DPSB08] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming. *3G Evolution: HSPA and LTE for Mobile Broadband*. Elsevier Academic Press, ISBN 9780123745385, LCCN 2008931278, 2nd edition, 2008.

[DRVV11] A. Drozdy, Rakos, Z. Vincze, and C. Vulkan. Adaptive VoIP Multiplexing in LTE Backhaul. In *IEEE 73rd Vehicular Technology Conference*, Yokohama, Japan, 15-18 May 2011.

[ELA11] G. A. Elkheir, A. S. Lioumpas, and A. Alexiou. Energy Efficient AF Relaying Under Error Performance Constraints with Application to M2M

Networks. In *IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications*, pages 56–60, Toronto, ON, Canada, 11-14 Sept. 2011.

[fD11] EXALTED Expanding LTE for Devices. WP2 - Business Models, Use cases and Technical Requirements. Technical Report FP7 Contract Number: 258512, European Commission Information Society and Media, 31 May 2011.

[Gol05] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 40 West 20th Street, New York, NY 10011-4211, USA, ISBN 9780521837163, LCCN 0521837162, 2005.

[GRG13] C. Gueguen, A. Rachedi, and M. Guizani. Incentive Scheduler Algorithm for Cooperation and Coverage Extension in Wireless Networks. *IEEE Transactions on Vehicular Technology*, 62(2):797–808, IEEE Vehicular Technology Society, ISSN 0018-9545, Feb. 2013.

[Gud91] M. Gudmundson. Correlation Model for Shadow Fading in Mobile Radio Systems. *Electronics Letters*, 27(23):2145–2146, Institution of Engineering and Technology, ISSN 0013-5194, 7 Nov. 1991.

[HGSO06] Jung Ha Hong, O. Gusak, K. Sohraby, and N. Oliver. Performance Analysis of Packet Encapsulation and Aggregation. In *14th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, pages 137–146, Monterey, CA, USA, 11-14 Sept. 2006.

[HHL86] H. Heffes, N. J. Holmdel, and D. Lucantoni. A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance. *IEEE Journal on Selected Areas in Communications*, 4(6):856–868, ISSN 0733-8716, Sept. 1986.

[HjNT+10] Qiong Huang, Ming jing Ni, Lun Tang, Rong Chai, and Qian bin Chen. Relay Protocol Improvement and Frame Structure Design Base on Overhearing Mechanism and Physical Network Coding. In *IEEE Youth Conference on Information Computing and Telecommunications*, pages 319–322, Beijing, China, 28-30 Nov. 2010.

[HKB11] E. Hossain, Dong In Kim, and V. K. Bhargava. *Cooperative Cellular Wireless Networks*. Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, United Kingdom, ISBN 9780521767125, LCCN 2010048066, 2011.

[HT05] G. Horváth and M. Telek. Analysis of a BMAP/D/1-Timer Multiplexer. *Electronic Notes in Theoretical Computer Science*, 128(4):25–44, Elsevier Science Publishers, Amsterdam, The Netherlands, ISSN 1571-0661, Apr. 2005.

[HT09] H. Holma and A. Toskala. *LTE for UMTS: OFDMA and SC-FDMA Based Radio Access*. John Wiley and Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, ISBN 9780470994016, LCCN 2008052792, 2009.

[HT11] H. Holma and A. Toskala. *LTE for UMTS Evolution to LTE-Advanced*. John Wiley and Sons, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, ISBN 9780470660003, LCCN 2010050375, 2nd edition, 2011.

[HTV03] G. Horváth, M. Telek, and C. Vulkán. AAL2 Multiplexing Delay Calculation in UTRAN. In *11th Microcoll Conference*, Budapest, Hungary, 10-11 Sept. 2003.

[HV06] G. Horváth and C. Vulkán. Queueing Model of the AAL2 Multiplexer in UTRAN. In *13th International Conference on Analytical and Stochastic Modelling Techniques and Applications*, Bonn, Germany, 28-31 May 2006.

[ID08] R. Irmer and F. Diehm. On Coverage and Capacity of Relaying in LTE-Advanced in Example Deployments. In *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, Cannes, France, 15-18 Sept. 2008.

[ITN10] M. Iwamura, H. Takahashi, and S. Nagata. Relay Technology in LTE-Advanced. *NTT DOCOMO Technical Journal*, 12(2):29–36, Sept. 2010.

[IWR10] J.C. Ikuno, M. Wrulich, and M. Rupp. System Level Simulation of LTE Networks. In *IEEE 71st Vehicular Technology Conference*, Taipei, Taiwan, 16-19 May 2010.

[JSC11] M. Jain, G. C. Sharma, and S. Chakrawarti. Performance Evaluation of an ATM Adaptation Layer 2 Multiplexer with Buffer. *Journal of Management and Information Technology*, 3(1):130–142, Guru Nanak Institute of Management, New Delhi, India, ISSN 0975-5187, 2011.

[KO13] Toshihito Kudo and Tomoaki Ohtsuki. Cell Range Expansion Using Distributed Q-learning in Heterogeneous Networks. *EURASIP Journal on Wireless Communications and Networking*, 6(6-3S):1029–1036, Springer, Cham, Switzerland, ISSN 1687-1499, 4 Mar. 2013.

[KY08] Ho Chee Keong and M. R. Yuce. Low Data Rate Ultra Wideband ECG Monitoring System. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3413–3416, Vancouver, BC, Canada, 20-25 Aug. 2008.

[KYMP12] N. Krishnan, R. D. Yates, N. B. Mandayam, and J. S. Panchal. Bandwidth Sharing for Relaying in Cellular Systems. *IEEE Transactions on Wireless Communications*, 11(1):117–129, IEEE Communications Society, ISSN 1536-1276, Jan. 2012.

[Law04] G. Lawton. Machine-to-Machine Technology Gears up for Growth. *Computer*, 37(9):12–15, IEEE Computer Society, ISSN 0018-9162, Sept. 2004.

[LTCH09] Shang-Wen Luan, Jen-Hao Teng, Shun-Yu Chan, and Lain-Chyr Hwang. Development of a Smart Power Meter for AMI Based on ZigBee Communication. In *International Conference on Power Electronics and Drive Systems*, pages 661–665, Taipei, Taiwan, 2-5 Nov. 2009.

[Lte13] LteWorld. LTE Operators. `http://www.lteworld.org/operator`, last accessed on 11 Sept. 2013.

[Mar02] J. Marconi. E-Health: Navigating the Internet for Health Information Healthcare, May 2002. Advocacy White Paper, Healthcare Information and Management Systems Society.

[Mar11] S. N. K. Marwat. Bandwidth and QoS Aware LTE Uplink Scheduler. In *Master Thesis*. ComNets - University of Bremen, Oct. 2011.

[McYZGTG12] S. N. K. Marwat, T. Weerawardane cand Y. Zaki, C. Goerg, and A. Timm-Giel. Performance Evaluation of Bandwidth and QoS Aware LTE Uplink Scheduler. In *10th International Conference on Wired/Wireless Internet Communications*, pages 298–306, Santorini, Greece, 6-8 June 2012.

[MDL+14] S. N. K. Marwat, Yangyang Dong, Xi Li, Y. Zaki, and C. Goerg. Novel Schemes for Component Carrier Selection and Radio Resource Allocation in LTE-Advanced Uplink. In *6th International Conference on Mobile Networks and Management*, Wuerzburg, Germany, 22-24 Sept. 2014.

[MEJA91] P. E. Mogensen, P. Eggers, C. Jensen, and J. B. Andersen. Urban Area Radio Propagation Measurements at 955 and 1845 MHz for Small and Micro Cells. In *Global Telecommunications Conference*, pages 1297–1302 vol.2, Phoenix, AZ, USA, 2-5 Dec. 1991.

[MM10] M. Menth and S. Mühleck. Packet Waiting Time for Multiplexed Periodic On/Off Streams in the Presence of Overbooking. *International Journal of Communication Networks and Distributed Systems*, 4(2):207–229, Inderscience Publishers, ISSN 1754-3916, 2010.

[MPKM08] G. Monghal, K. I. Pedersen, I. Z. Kovacs, and P. E. Mogensen. QoS Oriented Time and Frequency Domain Packet Schedulers for the UTRAN Long Term Evolution. In *IEEE 67th Vehicular Technology Conference*, pages 2532–2536, Singapore, 11-14 May 2008.

[MPZ+13] S. N. K. Marwat, T. Pötsch, Y. Zaki, T. Weerawardane, and C. Goerg. Addressing the Challenges of E-Healthcare in Future Mobile Networks. In *19th EUNICE/IFIP WG 6.6 International Workshop*, pages 90–99, Chemnitz, Germany, 28-30 Aug. 2013.

[MWZ+14] S. N. K. Marwat, T. Weerawardane, Y. Zaki, C. Goerg, and A. Timm-Giel. Analysis of Radio Resource Allocation in LTE Uplink. *Wireless Personal Communications*, 79(3):2305–2322, Springer, New York, USA, ISSN 0929-6212, Dec. 2014.

[MZC+14] S. N. K. Marwat, Y. Zaki, J. Chen, A. Timm-Giel, and C. Goerg. A Novel Machine-to-Machine Traffic Multiplexing in LTE-A System using Wireless In-band Relaying. In *5th International Conference on Mobile Networks and Management*, pages 149–158, Cork, Ireland, 23-25 Sept. 2014.

[NH12] Tien-Dung Nguyen and Eui-Nam Huh. A Dynamic ID-Based Authentication Scheme for M2M Communication of Healthcare Systems. *The International Arab Journal of Information Technology*, 9(6):511–519, Zarqa University, Jordan, ISSN 2309-4524, Nov. 2012.

[NXW11] D. Niyato, Lu Xiao, and Ping Wang. Machine-to-Machine Communications for Home Energy Management System in Smart Grid. *IEEE Communications Magazine*, 2(3):53–59, IEEE Communications Society, ISSN 0163-6804, Apr. 2011.

[one14] oneM2M. . `http://www.onem2m.org`, last accessed on 9 Oct. 2014.

[PMZG13] T. Pötsch, S. N. K. Marwat, Y. Zaki, and C. Goerg. Influence of Future M2M Communication on the LTE System. In *6th Joint IFIP Wireless and Mobile Networking Conference*, Dubai, UAE, 23-25 Apr. 2013.

[Pok07] A. Pokhariyal. Downlink Frequency-Domain Adaptation and Scheduling - A Case Study Based on the UTRA Long Term Evolution. In *PhD Thesis*. the Faculty of Engineering, Science and Medicine, Aalborg University, Aalborg, Denkmark, Aug. 2007.

[Riv14] Riverbed. Riverbed application and network performance management solutions. `http://www.opnet.com`, last accessed on 14 Aug. 2014.

[RRRH09] A. A. Rasheed, S. Redana, B. Raaf, and J. Hamalainen. Uplink Resource Partitioning in Relay Enhanced LTE-Advanced Networks. In *20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1502–1506, Tokyo, Japan, 13-16 Sept. 2009.

[Sem14] Freescale Semiconductor. Long Term Evolution Protocol Overview. `https://www.freescale.com/files/wireless_comm/doc/white_paper/LTEPTCLOVWWP.pdf`, last accessed on 9 Oct. 2014.

[Sho04] 3GPP TSG RAN Future Evolution Work Shop. Compendium of Abstracts. Technical report, 3rd Generation Partnership Project, Toronto, Canada, Nov. 2004.

[SJL+12] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, and J. Wang. A First Look at Cellular Machine-to-Machine Traffic: Large Scale Measurement and Characterization. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement*

and Modeling of Computer Systems, pages 65–76, New York, NY, USA, June 2012.

[SRMG09] B. Scholz-Reiter, A. Mehrsai, and M. Görges. Handling Dynamics in Logistics - Adoption of Dynamic Behaviour and Reduction of Dynamic Effects. *Asian International Journal of Science and Technology in Production and Manufacturing Engineering*, 2(3):99–110, King Mongkut's University of Technology North Bangkok Press, ISSN 1906-151X, 2009.

[STB09] S. Sesia, I. Toufik, and M. Baker. *LTE, The UMTS Long Term Evolution: From Theory to Practice*. John Wiley and Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, ISBN 9780470697160, LCCN 2008041823, 2009.

[TFP+10] O. Teyeb, F. Frederiksen, Vinh Van Phan, B. Raaf, and S. Redana. User Multiplexing in Relay Enhanced LTE-Advanced Networks. In *IEEE 71st Vehicular Technology Conference*, Taipei, Taiwan, 16-19 May 2010.

[TV05] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, ISBN 9780521845274, LCCN 2006272166, 2005.

[Vod14] Vodafone. Machine-to-Machine (M2M) Transport and Logistics. `http://m2m.vodafone.com/discover-m2m/industries/transport-and-logistics`, last accessed on 7 Oct. 2014.

[Voj08] A. Vojdani. Smart Integration. *IEEE Power and Energy Magazine*, 6(6):71–79, IEEE Power and Energy Society, ISSN 1540-7977, Nov.-Dec. 2008.

[VWHS09] V. Venkatkumar, T. Wirth, T. Haustein, and E. Schulz. Relaying in Long Term Evolution: Indoor Full Frequency Reuse. In *European Wireless Conference*, pages 298–302, Aalborg, Denmark, 17-20 May 2009.

[WJR12] Wu Wen, Maozu Jin, and Peiyu Ren. TD-LTE Uplink Power Control Research Based on Users Location. *Applied Mathematics and Information Sciences*, 6(6-3S):1029–1036, Natural Sciences Publishing, ISSN 1935-0090, 26 Feb. 2012.

[WPSM04] D. S. Watson, M. A. Piette, O. Sezgen, and N. Motegi. Machine to Machine (M2M) Technology in Demand Responsive Commercial Buildings. In *ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, USA, Aug. 2004.

[WRP11] Hua Wang, C. Rosa, and K. I. Pedersen. Uplink Component Carrier Selection for LTE-Advanced Systems with Carrier Aggregation. In *IEEE International Conference on Communications*, Kyoto, Japan, 5-9 June, 2011.

[WTJ+11] Geng Wu, S. Talwar, K. Johnsson, N. Himayat, and K. D. Johnson. M2M: From Mobile to Embedded Internet. *IEEE Communications Magazine*, 49(4):36–43, IEEE Communications Society, ISSN 0163-6804, Apr. 2011.

[WVH+09] T. Wirth, V. Venkatkumar, T. Haustein, E. Schulz, and R. Halfmann. LTE-Advanced Relaying for Outdoor Range Extension. In *IEEE 70th Vehicular Technology Conference*, Anchorage, AK, USA, 20-23 Sept. 2009.

[XRG+06] Weimin Xiao, R. Ratasuk, A. Ghosh, R. Love, Sun Yakun, and R. Nory. Uplink Power Control, Interference Coordination and Resource Allocation for 3GPP E-UTRA. In *IEEE 64th Vehicular Technology Conference*, Montreal, Quebec, Canada, 25-28 Sept. 2006.

[Zak12] Y. Zaki. *LTE Optimization and Mobile Network Virtualization*. Springer, ISBN 9783658008079, LCCN 2012953090, 2012.

[ZMS+08] Naizheng Zheng, P. H. Michaelsen, J. Steiner, C. Rosa, and J. Wigard. Antenna Tilt and Interaction with Open Loop Power Control in Homogeneous Uplink LTE Networks. In *IEEE International Symposium on Wireless Communication Systems*, pages 693–697, Reykjavík, Iceland, 21-24 Oct. 2008.

[ZWGTG11a] Y. Zaki, T. Weerawardane, C. Goerg, and A. Timm-Giel. Multi-QoS-Aware Fair Scheduling for LTE. In *IEEE 73rd Vehicular Technology Conference*, Yokohama, Japan, 15-18 May 2011.

[ZWGTG11b] Y. Zaki, T. Weerawardane, C. Goerg, and A. Timm-Giel. Long Term Evolution (LTE) Model Development within OPNET Simulation Environment. In *OPNET Workshop*, Washington, D.C., USA, Aug. 29-Sept. 1, 2011.

[ZZW+11] Y. Zaki, N. Zahariev, T. Weerawardane, C. Goerg, and A. Timm-Giel. Optimized Service Aware LTE MAC Scheduler: Design, Implementation and Performance Evaluation. In *OPNET Workshop*, Washington, D.C., USA, Aug. 29-Sept. 1, 2011.